

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 20:05:40 ; Search time 571 Seconds
(without alignments)
10211.813 Million cell updates/sec

Title: US-10-089-543-2_COPY_449_1433

Perfect score: 985

Sequence: 1 gatgtaccagtggtggag.....tgagagaatccttcacatc 985

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	985	100.0	1433	6	ABK15699 Cotton fi
2	107.4	10.9	8056	8	ABZ10246 Haematopo
3	105.8	10.7	8056	8	ABZ10100 Haematopo
C	104.8	10.6	158001	12	AD117884 Human pho
5	104.4	10.6	4985	6	ABQ75107 Anopheles
6	104.4	10.6	4985	10	ACF79720 Mosquito
7	102.4	10.4	8222	8	ACF62816 Colon can
8	102.4	10.4	11222	10	ADB54318 Pretreat
C	101.6	10.3	8056	8	ABZ10246 Haematopo
10	100.8	10.2	8222	8	ACF62794 Colon can
11	100.8	10.2	11222	10	ADB54190 Pretreat
C	100.0	10.2	8056	8	ABZ10100 Haematopo
13	96.2	9.8	15548	6	ABL34155 Human imm
14	95.4	9.7	6109	6	ABL32326 Human imm
15	95.4	9.7	6109	6	AAS61077 Human gen
16	94.6	9.6	6963	6	ABL32979 Human imm
17	94.4	9.6	419	8	ABX46069 Bovine ES
C	94.4	9.6	5286	13	ADS89278 Oligonuc
18	94.4	9.6	5286	13	ADS89278 Oligonuc
C	94.4	9.6	5286	13	ADS89552 Oligonuc
20	93.6	9.5	9642	6	ABL32357 Human imm

21	91.8	9.3	6216	6	ABK39932 Human che
22	91.8	9.3	6216	6	ABL70139 Chemical
23	91.8	9.3	7167	6	ABL32400 Human imm
24	91.2	9.3	6048	6	ABQ67002 Human ang
25	91.2	9.3	12237	6	ABL34358 Human imm
26	90.8	9.2	7814	4	AAS46530 Tumour su
27	90.6	9.2	6419	6	ABL32267 Human imm
C	90.2	9.2	778	6	ABQ15588 Oligonuc
28	90.2	9.2	778	6	ABQ15589 Oligonuc
29	90.2	9.2	6175	6	ABL33307 Human imm
30	90.2	9.2	6175	6	ABL33307 Human imm
31	90	9.1	18154	6	ABL32254 Human imm
32	89.6	9.1	6145	6	ABL32972 Human imm
C	89.2	9.1	9789	2	AAT41852 cDNA enco
34	89	9.0	3683	8	ABZ10199 Haematopo
C	88.6	9.0	6644	2	AAX33181 Base sequ
36	88.6	9.0	7372	2	AAX33182 Base sequ
C	88.6	9.0	7797	2	AAX33180 Cowpox vi
37	88.6	9.0	7797	2	AAX33180 Cowpox vi
C	88.4	9.0	7996	2	AAX33184 Base sequ
39	88.4	9.0	3738	3	AAA70178 Plasmodu
C	88.4	9.0	5768	6	ABK31192 Signal tr
41	88.4	9.0	5768	6	ABK31192 Signal tr
42	88.4	9.0	5768	6	ABL70517 Chemical
43	88.4	9.0	6286	4	AAS61105 Human gen
44	88.4	9.0	61020	4	AAS46591 Tumour su
45	88.2	9.0	6292	4	AAS46787 Tumour su
					AAS46735 Tumour su

ALIGNMENTS

RESULT 1
ABK15699
ID ABK15699 standard; DNA; 1433 BP.
XX
AC ABK15699;
XX
DT 21-MAY-2002 (first entry)
XX
DE Cotton fibre-specific beta tubulin, CFTUB2, promoter fragment.
XX
KW Cotton; ds; CFTUB2; fibre-specific beta tubulin; transgenic. promoter;
KW plant; anthocyanin gene; silk protein gene; cotton fibre strength;
KW polyhydroxybutyrate.
XX
OS Gossypium hirsutum.
XX
FH Key Location/Qualifiers
FT Promoter 449..1433
FT /*tag= a
FT /note= "Cotton fibre-specific promoter. This sequence is specifically claimed in claim 3"
XX
PN WO200210377-A1.
XX
PD 07-FEB-2002.
XX
PF 01-AUG-2000; 2000WO-SG000111.
XX
PR 01-AUG-2000; 2000WO-SG000111.
XX
PA (MOLE-) INST MOLECULAR AGROBIOLOGY.
XX
PI Cai L, Li X, Cheng N, Liu J;
XX
DR WPI; 2002-217120/27.
XX
PT New fiber-specific beta-tubulin promoter from cotton for controlling gene
PT expression in cotton fibers and creating transgenic plants, in particular
PT cotton plants, having altered fiber characteristics.
XX
PS Claim 2; Fig 2; 30pp; English.
XX
CC The invention relates to a promoter that is cotton fibre-specific,

comprising the promoter of the cotton beta-tubulin gene CFTUB2. The promoter is fibre-specific in cotton and controls specific gene expression at the transcriptional level in cotton fibres and is useful for improving cotton fibres to create new cotton varieties with high fibre quality and yield by gene manipulation. The promoter is useful for creating transgenic plants, in particular cotton having altered fibre characteristics, and permits selective expression of a transgene in the cotton fibre, permitting greater latitude in the types of transgenes employed. Examples of expression of desirable genes in cotton fibre, but not in other parts of the cotton plants, include anthocyanin genes for coloured cotton, silk protein genes from silk worm or spiders for increased strength of cotton fibre, and biosynthesis of polyhydroxybutyrate in cotton fibre for improved thermal properties and insulating characteristics. The promoter can improve cotton fibres to create new cotton varieties with higher fibre quality and yield. The present sequence is the CFTUB2 promoter fragment

XX SQ Sequence 1433 BP; 466 A; 285 C; 180 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 985; DB 6; Length 1433;
Best Local Similarity 100.0%; Pred. No. 4.3e-123;
Matches 985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGTGACCAAGTGGTGGAGAGAGAGGCTACCGATTTGGTCAAGTGGGACCAATTTTGA 60
DB |||||
QY 61 TTTTACCCTCCTCCAGTAGATTCCTAAATCTATGCAATTTATCTCATTTCAATTTATTT 120
DB |||||
QY 509 TTTTACCCTCCTCCAGTAGATTCCTAAATCTATGCAATTTATCTCATTTCAATTTATTT 568
DB |||||
QY 121 AATTATTTTATATTTTTCGATAAAATCTAATCTTTTACCTTTTATTTTAAAAAGAAATTT 180
DB |||||
QY 569 AATTATTTTATATTTTGGATATAAAATCTAATCTTTTACCTTTTATTTTAAAAAGATTT 628
DB |||||
QY 181 AATTATTTTATATTTTATATTTAGATAAAATCTAATCTTTTACCTTTTATTTTAAAAAG 240
DB |||||
QY 629 AATTATTTTATATTTTATATTTAGATAAAATCTAATCTTTTACCTTTTATTTTAAAAAG 688
DB |||||
QY 241 AATTCAATGCGTTTTCCTAATTTAGTTTAACTTCTATCTACTAATTAATAAATTCCTG 300
DB |||||
QY 689 AATTCAATGCGTTTTCCTAATTTAGTTTAACTTCTATCTACTAATTAATAAATTCCTG 748
DB |||||
QY 301 ATCGGATTTAGTGTGGTCAAAAGTCAAGTCAATGAATTTTGTGGAGAAAAATAAAAA 360
DB |||||
QY 749 ATCGGATTTAGTGTGGTCAAAAGTCAAGTCAATGAATTTTGTGGAGAAAAATAAAAA 808
DB |||||
QY 361 TTAACACATTTTTCGATTAATTTATATATATATATATATATATATATATATATATAT 420
DB |||||
QY 809 TTAACACATTTTTCGATTAATTTATATATATATATATATATATATATATATATATAT 868
DB |||||
QY 421 ATGTTGTCAATATATATTTTAAATTTAAATTTTCAGCACAAATTTACACTCTCATCATTA 480
DB |||||
QY 869 ATGTTGTCAATATATTTTAAATTTAAATTTTCAGCACAAATTTACACTCTCATCATTA 928
DB |||||
QY 481 AATTATCTTATACCAATTAATTTTGTGGAGCAATTTATTTTAACTCACCCTTC 540
DB |||||
QY 929 AATTATCTTATACCAATTAATTTTGTGGAGCAATTTATTTTAACTCACCCTTC 988
DB |||||
QY 541 CATTAATGCAATATATATTTTGTGGAGCAATTTATTTTAACTCACCCTTC 1048
DB |||||
QY 989 CATTAATGCAATATATATTTTGTGGAGCAATTTATTTTAACTCACCCTTC 1048
DB |||||
QY 601 TTAACCAATTTTGAAGTCTTAAATTTCTTAACTTATTTTACATTTTGTGGTCTGGGTCC 660
DB |||||
QY 1049 TTAACCAATTTTGAAGTCTTAAATTTCTTAACTTATTTTACATTTTGTGGTCTGGGTCC 1108
DB |||||
QY 661 ATCTGAAAGGCGCAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCT 720
DB |||||
QY 1109 ATCTGAAAGGCGCAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCTCAAGCTCT 1168
DB |||||
QY 721 AACTACATTTGTACATTTAAGCAATTTCCCAATTTCCCAATTTCCCAATTTCCCAATTT 780
DB |||||

Db 1169 AACTACATTTGTACAGTTTACTAAGCAAAATCCAAATTTCAAAAATTCATTTTCCAGGAAA 1228
QY 781 ACGAAACGTCGTTTACTAATACCGACCTAAACCCAGCTCAACCTGCGTCAATTTAACGGAA 840
Db 1229 ACGAAACGTCGTTTACTAATACCGACCTAAACCCAGCTCAACCTGCGTCAATTTAACGGAA 1288
QY 841 ATCTTTTAACTCTCTATATATAACCAAAACCACTCTCATCAATTTCCCATATAAAGA 900
Db 1289 ATCTTTTAACTCTCTATATATAACCAAAACCACTCTCATCAATTTCCCATATAAAGA 1348
QY 901 ATTTCCGGAATTTCTTATTTCTTTTATATTTTCTCTCCAAATTTCCCTCACTTTTCCGGA 960
Db 1349 ATTTCCGGAATTTCTTATTTCTTTTATATTTTCTCTCCAAATTTCCCTCACTTTTCCGGA 1408
QY 961 GAAATGAGAGAAATCCCTTCACATC 985
Db 1409 GAAATGAGAGAAATCCCTTCACATC 1433

RESULT 2
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX AC ABZ10246;
XX DT 16-JAN-2003 (first entry)
XX Haematopoietic cell proliferation disorder related DNA sequence #386.
DE Human; haematopoietic cell proliferation disorder; cytoskeletal;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
OS Homo sapiens.
XX PN WO200277272-A2.
XX PD 03-OCT-2002.
XX PF 26-MAR-2002; 2002WO-EP003401.
XX PR 26-MAR-2001; 2001US-0278333P.
XX (EPIG-) EPIGENOMICS AG.
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
XX Olek A, Piepenbrock C, Adorian P, Grabs G, Lesche R, Leu E;
XX Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
XX Schwobe I, Ziebarth H;
XX WPI; 2003-018942/01.
XX Detecting and differentiating between hematopoietic cell proliferative
XX disorders, comprises contacting a target nucleic acid with a reagent that
XX distinguishes between methylated and non-methylated CpG dinucleotides.
XX Claim 28; SEQ ID NO 386; 117pp; English.
XX The present invention describes a method for detecting and
XX differentiating between haematopoietic cell proliferative disorders
XX associated with at least 1 gene and/or their regulatory regions in a
XX subject. The method comprises contacting a target nucleic acid in a
XX biological sample obtained from the subject with at least 1 reagent,
XX which distinguishes between methylated and non-methylated CpG
XX dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
XX represent specifically claimed nucleotide sequences from the present
XX invention. Oligonucleotides from the present invention can be used for
XX differentiating between healthy haematopoietic cells and proliferative
XX disorder haematopoietic cells; for differentiating between acute
XX lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
XX determining the cytosine methylation state and/or single nucleotide
XX polymorphisms (SNPs) of haematopoietic cell proliferation disorder
XX related sequences and their complements; and as primers for the

CC	amplification of haematopoietic cell proliferation disorder related DNA sequences. The nucleotide sequences from the present invention can also be used for detecting a predisposition to, differentiation between subclasses, diagnosis, prognosis, treatment and/or monitoring of haematopoietic cell proliferative disorders. The present method enables a highly specific classification of haematopoietic cell proliferative disorders allowing for improved and informed treatment of patients
XX	
SQ	Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
	Query Match 10.9%; Score 107.4; DB 8; Length 8056;
	Best Local Similarity 51.1%; Pred No. 4.8e-06;
	Matches 305; Conservative 0; Mismatches 286; Indels 6; Gaps 2;
Qy	52 AATTTTATTATTACCTCCTCGCTAGATTCGTAATACTATTGCATTTATCTCATTTTCAT 111
Db	55 AAATTTTTTTTTTATTTTTTTTTTAAATTTTTTAAATTTTTTATATATTTTTTATATATAAT 114
Qy	112 TATTATTTAATATTTTATATATTTTGGATAAAAAATCTCAATACTTTTACTTTTTTTTAA 171
Db	115 TATTATTTTAAATTTTTTATATTTTTTTTATTTATTTTATTTTATTTTATTTTATTTT-- 172
Qy	172 AAAGAATTTTAAATTTATTTATTTATTTTAGATAAAAAATCTCAAATCTTTACTTTTTT 231
Db	173 TATTATTTTTTTTAAATTTATTTAAATTTTAAATTTATTTATTTTAAAATAATTTTAAA 232
Qy	232 TTTAAAAGAATTTCAATTCGGTGTTTTTTCTTAAATTTAGTTTTTAAATCTTACTAATTATA 291
Db	233 ATTTTAAAAAATTTAAAAAATTTTAAAAATTTTAAATTTTATTTTAAAAAAAATTTTTT 292
Qy	292 AAAATCTGATCGGAATTAGTGTGTCAAAGTCAAGTCACATGAAATTTGTTGGAGAAA 351
Db	293 ATTTTTTTTTAAAAATTTATTTTTTAAAAAATAATTAATAAAAAATAAATAAAAAATAATA 352
Qy	352 AAATABAAAAATTAACACATTTTTTCGATTAATTTTATTTATATATATAATAATAAACACAT 411
Db	353 ATTAAAAAATAAAAAAATAAAAAATGAAATGAAATATTTTAAATTTTAAATTTTAAATATAT 412
Qy	412 TTTTATTAAATGT----TGTCAATAATTTTTTTTAAATTTTAAATTTTCAGCACACAATTAC 467
Db	413 ATTTATTAAATTTATTTTATTTTTTTTAAAAATAATAAAAAAATAATTTTAAAAAATAA 472
Qy	468 ACTCTCATCATTAATTTTAAATTTTATTTACCATAAATTAATAATTTGTGAGGACAATTTT 527
Db	473 AATATATTTAAATAAAAAAATAAAAAATTTATAATTTTATAATTTATAATTAATTAATATAT 532
Qy	528 TAATCTCACCCTCCAATTAATGCATATTTATTAATTTTGTTCGATCTCTTATTTTCATC 587
Db	533 TAAAAATTAATAATTTTATATAATTTATTAATGTTTTTATTAATAAATTTATATTTTAAATA 592
Qy	588 CTAAACATTTAATCATTTAACCCCAATTTTGAACGTGTTATTAATTTCTTAACTTATTCAC 644
Db	593 ATAATATTTTAAAAAATAATTTATATAAATTTTAAATTTTTTAAAAAATAATTTTAA 649
RESULT 3	
ABZ10100	
ID	ABZ10100 standard; DNA; 8056 BP.
XX	
AC	ABZ10100;
XX	
DT	16-JAN-2003 (first entry)
XX	
DE	Haematopoietic cell proliferation disorder related DNA sequence #240.
XX	
KW	Human; haematopoietic cell proliferation disorder; cytostatic;
KW	gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW	cytosine methylation state; gene; ds.
OS	Homo sapiens.
XX	
PX	WO200277272-A2.
XX	

PD		03-OCT-2002.
XX		
PF		26-MAR-2002; 2002WO-EP003401.
PP		
PR		26-MAR-2001; 2001US-0278333P.
PX		(EPIG-) EPIGENOMICS AG.
PA		Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
XX		Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI		Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI		Schwöpe I, Ziebarth H;
XX		
DR		WPI; 2003-018942/01.
XX		Detecting and differentiating between hematopoietic cell proliferative
PT		disorders, comprises contacting a target nucleic acid with a reagent that
PT		distinguishes between methylated and non-methylated CpG dinucleotides.
XX		
PS		Claim 28; SEQ ID NO 240; 117pp; English.
XX		The present invention describes a method for detecting and
CC		differentiating between haematopoietic cell proliferative disorders
CC		associated with at least 1 gene and/or their regulatory regions in a
CC		subject. The method comprises contacting a target nucleic acid in a
CC		biological sample obtained from the subject with at least 1 reagent,
CC		which distinguishes between methylated and non-methylated CpG
CC		dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC		represent specifically claimed nucleotide sequences from the present
CC		invention. Oligonucleotides from the present invention can be used: for
CC		differentiating between healthy haematopoietic cells and proliferative
CC		disorder haematopoietic cells; for differentiating between acute
CC		lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC		determining the cytosine methylation state and/or single nucleotide
CC		polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC		related sequences and their complements; and as primers for the
CC		amplification of haematopoietic cell proliferation disorder related DNA
CC		sequences. The nucleotide sequences from the present invention can also
CC		be used for detecting a predisposition to, differentiation between
CC		subclases, diagnosis, prognosis, treatment and/or monitoring of
CC		haematopoietic cell proliferative disorders. The present method enables a
CC		highly specific classification of haematopoietic cell proliferative
XX		disorders allowing for improved and informed treatment of patients
XX		
SQ		Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;
	Query Match	10.7%; Score 105.8; DB 8; Length 8056;
	Best Local Similarity	50.9%; Pred. No. 7.9e-06;
	Matches 304; Conservative	0; Mismatches 287; Indels 6; Gaps 2
QY	52	AATTTTATTATTACCTCCTCGCTAGATTCGTAAATACTATTGCATTTATCTCATTTCAAT 111
DB		
QY	55	AAAATTTTTTTTTTAATTTTTTTTTTTAAATTTTTTAAATTTTATATTTATTTTATATATTAAT 114
DB		
QY	112	TATTATTATTAATTTTATATATTTATGGATAAAAAATTTCTAATACTTTTACTTTTTTTTAA 171
DB		
QY	115	TATTATTTTTAAATTTTTTATATTTTTTTTTTTATTTATTTTTTTTATTTATTTT-- 172
DB		
QY	172	AAAGAATTTATTTAATTTATTTATATTTTAGATAAAAAATTCATAACTTTTACTTTTTT 231
DB		
QY	173	TATTATTTTTTTAAATTTAATTTTAAATTTTAAATTTTATTTTAAATAAAAAATTTTAA 232
DB		
QY	232	TTTAAAGAATTTCAATTGCGTTTTTTTCTTAAATTAGTTTTAACTTATCTATACTAATATA 291
DB		
QY	233	ATTTTAAATAATTTAAATAATTTTAAATTTTAAATTTTATTTTAAATAAAAAATTTT 292
DB		
QY	292	AAAAATTTCTGATCGGATTAGTGCTGGTGTCAAAGTCAAGTCAATGAATTTTGTGTGGAGAA 351
DB		
QY	293	ATTTTTTTTAAATAATTTTAAAAATAATTAATAAAAAATAAAAAATAATAATAATA 352
DB		
QY	352	AAATAAAAAATTAACAACATTTTTTCGATTAAATTTATATATATATATATATAATAAACACAT 411
DB		
QY	353	ATTAAAAAAAAAAAAAAAAACGAAATAAAAAATTTTAAATTTAAAAATTTAAATTTAAATAAT 412
DB		

CC for treating diseases or conditions associated with phosphotyrosyl
 CC phosphatase activator, preferably hyperproliferative disorder or
 CC developmental disorder. The compound can also be used as prophylaxis,
 CC e.g. to prevent or delay infection, inflammation or tumour formation. The
 CC present sequence is the human PTPA gene which is a target for the
 CC antisense oligonucleotides of the invention.

XX
 SQ Sequence 158001 BP; 34764 A; 40497 C; 41696 G; 40433 T; 0 U; 611 Other;

Query Match 10.6%; Score 104.8; DB 12; Length 158001;
 Best Local Similarity 49.7%; Pred. No. 7.7e-06;
 Matches 297; Conservative 0; Mismatches 297; Indels 4; Gaps 1;

QY 52 AATTTTATTTTATTTTCTCTGCTAGATTCGTAATACTATTGCAATTTATCTCATTTTCAT 111
 DB 22812 ATTTTTTAT 22753

QY 112 TATTTTATTAATTTTAT 171
 DB 22752 TATATATTTGATATATTTTTTATATATATATATATATATATATATATATATATAT 22693

QY 172 AAAGAAATTTTATTAATTTTATATATATATATATATATATATATATATATATATAT 227
 DB 22692 AT 22633

QY 228 TTTTTTAAAGAAATTTCAATTCGCTTTTCTTAAATTTAGTTTAAATTTCTATATCTAAT 287
 DB 22632 TTTAT 22573

QY 288 TATAAAATTTCTGATCGGATAGTGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGA 347
 DB 22572 TAT 22513

QY 348 GAAAAATTAATAATTAACACATTTTTCGATTAATTTATATATATATATATATATATAT 407
 DB 22512 ATATTTTAT 22453

QY 408 ACATTTTATTTAAATTTGTCGAATAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAC 467
 DB 22452 TATTATATATTTTAT 22393

QY 468 ACTCTCATCATTAATTTAACTTTATACCATAATTAATAATTTGTCGAGGACATTTATTTT 527
 DB 22392 AAT 22333

QY 528 TAATCTCACCTCCATTAATGCAATATATTAATTTTGTTCGATACCTTCTTATTTTCACCTC 587
 DB 22332 TTAT 22273

QY 588 CTAAATTAATCATTAACCCCAATTTTGAACCTGTTAATAATTTCTTAACTTTATTCACCTA 645
 DB 22272 ATATATTTAT 22215

RESULT 5
 ABQ75107
 ID ABQ75107 standard; cDNA; 4985 BP.
 XX
 AC ABQ75107;
 XX
 DT 01-NOV-2002 (first entry)
 XX
 DE Anopheles gambiae odourant receptor 2 genomic DNA SEQ ID NO:10.
 XX
 KW Anopheles gambiae; mosquito; olfactory gene; arrestin 1; pest control;
 XX
 KW odourant receptor; olfaction; gene; ds.
 XX
 OS Anopheles gambiae.
 XX
 XX WO200259274-A2.
 XX
 PD 01-AUG-2002.
 XX

ADL17884/c
 ID ADL17884 standard; cDNA; 158001 BP.
 XX
 AC ADL17884;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human phosphotyrosyl phosphatase activator, PTPA, gene.
 XX
 KW Human; ds; antisense; phosphotyrosyl phosphatase activator; PTPA;
 KW hyperproliferative disorder; developmental disorder; infection;
 KW inflammation; tumour; gene.
 XX
 OS Homo sapiens.
 XX
 PN US2004023906-A1.
 XX
 PD 05-FEB-2004.
 XX
 PF 01-AUG-2002; 2002US-00211179.
 XX
 PR 01-AUG-2002; 2002US-00211179.
 XX
 XX (ISIS-) ISIS PHARM INC.
 XX
 PI Dean NM, Dobie KW;
 XX
 DR WPI; 2004-132607/13.
 XX
 PT New antisense compound targeted to a nucleic acid molecule encoding
 PT phosphotyrosyl phosphatase activator, for modulating expression of
 PT phosphotyrosyl phosphatase activator or treating hyperproliferative
 PT disorders.

Example 15; SEQ ID NO 11; 131pp; English.

The invention relates to a compound 8-80 nucleobases in length targeted to a nucleic acid molecule encoding phosphotyrosyl phosphatase activator (PTPA), that specifically hybridises with the nucleic acid molecule encoding phosphotyrosyl phosphatase activator and inhibits the expression of phosphotyrosyl phosphatase activator, i.e. an antisense oligonucleotide. Also included are a composition comprising the compound and a pharmaceutical carrier or diluent, a method of inhibiting the expression of phosphotyrosyl phosphatase activator in cells or tissues, a method of treating an animal having a disease or condition associated with phosphotyrosyl phosphatase activator and a method of screening for an antisense compound. The disease or condition is a hyperproliferative disorder or developmental disorder. The compound, particularly the antisense oligonucleotide is useful in modulating the function of nucleic acid molecules encoding phosphotyrosyl phosphatase activator. The antisense compound can also be used as research tools and diagnostics. It can also be used as tools in differential and/or combinatorial analyses to elucidate expression patterns of a portion or the entire complement of genes expressed within cells and tissues. The compound can also be used

SQ	Sequence	8222 BP; 2010 A; 0 C; 1769 G; 4443 T; 0 U; 0 Other;
XX	Query Match	10.4%; Score 102.4; DB 8; Length 8222;
XX	Best Local Similarity	51.5%; Pred. No. 2.2e-05;
XX	Matches	310; Conservative 0; Mismatches 286; Indels 6; Gaps 3;
QY	53	ATTTTATTTTACCTCGCTAGATTTCGTAATAACTATGTCATTTCTCATTTCAATT 112
DB	7312	AGTTTATTTTATTTATGTAATTTTATTTTATTTTATTTATTTAGTATTTAATTTATTTATTT 7371
QY	113	ATTTATTTTAAATTTT--TTATATTTTTCGATAAAATTTCTAATCTTTTACTTTTAA 170
DB	7372	ATTTATTTTATTTGTTATTTAGTATTTATTTATTTATTTATTTGTTATTTTAAATTTA 7431
QY	171	AAAAGAAATTTATTTATTTATTTATTTATTTAGATAAAATTTCTAATCTTTTACTTTTT 230
DB	7432	TTTATTTATTTATTTTGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7491
QY	231	TTTTAAAAAGAAATTCGAATTCGTTTCTTAATTTAGTTTAAATTTCTATATCTAATTTAT 290
DB	7492	TATTTAATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 7551
QY	291	AAAAATTTCTGATCGGATTTAGTGGTCAAGTCAAGTCACATGAATTTTGTGGAGAA 350
DB	7552	TTTATTTATTTTGTGTTATTTAGGTAATTTATTTATTTATTTATTTATTTATTTATTTAA 7611
QY	351	AAAAATTTAAACACACATTTTTCGATTAATTTATTTATATATATATATATAATAATAACACA 410
DB	7612	AAATTTATTTAGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7671
QY	411	TTTTTATTTAAAGTGTGCAATATTTTCTTAATTTAGTTTAAATTTCTAATCTTTTACTTTT 468
DB	7672	TTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7731
QY	469	CTCTCATCTAAATTTAATCTTATTTACCATAATTTAAATTTGTCGAGACAAATTTATTTT 528
DB	7732	ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7791
QY	529	AAATCTCACCTCCATTAATGATATTTATTAATTTTGTTCGA--TACTTCTTATTTTCACT 586
DB	7792	TATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7851

XX	(SPIG-) EPIGENOMICS AG.
XX	Adorjan P, Burger M, Maier S, Nimmrich I, Becker E, Lesche R;
XX	Rujan T, Schmitt A;
XX	WPI; 2003-731620/69.
XX	Detecting and differentiating between colon cell proliferative disorders
XX	associated with a gene or its regulatory regions comprises contacting a
XX	target nucleic acid in a biological sample obtained from the subject with
XX	a reagent.
XX	Claim 32; SEQ ID NO 374; 74pp; English.
XX	The invention relates to a novel method for detecting and differentiating
XX	between colon cell proliferative disorders associated with at least one
XX	gene or its regulatory regions. The method comprises contacting a target
XX	nucleic acid in a biological sample obtained from the subject with at
XX	least one reagent or a series of reagents, where the reagent or series of
XX	reagents, distinguishes between methylated and non methylated CpG
XX	dinucleotides within the target nucleic acid. The molecules of the
XX	invention demonstrate cytosine methylation state or single nucleotide
XX	polymorphisms. The current sequence is that of the pretreated genomic DNA
XX	region of the invention. This sequence is not shown within the
XX	specification but is taken from Wipoweb.
SQ	Sequence 11222 BP; 2767 A; 0 C; 2498 G; 5957 T; 0 U; 0 Other;
	Query Match 10.4%; Score 102.4; DB 10; Length 11222;
	Best Local Similarity 51.5%; Pred. No. 2.2e-05;
	Matches 310; Conservative 0; Mismatches 286; Indels 6; Gaps 3;
QY	53 ATTTTATTTTACCTCGCTAGATTTCGTAATAACTATGTCATTTCTCATTTCAATT 112
DB	7312 AGTTTATTTTATTTATGTAATTTTATTTTATTTTATTTAGTATTTAATTTATTTATTT 7371
QY	113 ATTTATTTAAATTT--TTATATTTTTCGATAAAATTTCTAATCTTTTACTTTTAA 170
DB	7372 ATTTATTTTATTTGTTATTTAGTATTTATTTATTTATTTATTTATTTTGTATTTTAAATTTA 7431
QY	171 AAAAGAAATTTATTTATTTATTTATTTATTTAGATAAAATTTCTAATCTTTTACTTTTT 230
DB	7432 TTTTATTTATTTATTTTGTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7491
QY	231 TTTTAAAAAGAAATTCGAATTCGTTTCTTAATTTAGTTTAAATTTCTATATCTAATTTAT 290
DB	7492 TATTTAATATTTTATTTATTTTATTTATTTATTTATTTATTTTAAATTTATTTATTTATTTA 7551
QY	291 AAAAAATTTCTGATCGGATTTAGTGGTCAAGTCAAGTCACATGAATTTTGTGGAGAA 350
DB	7552 TTTTATTTATTTTGTGTTATTTAGGTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 7611
QY	351 AAAATAAAATTTAAACACATTTTTCGATTAATTTATTTATATATATATAATAATAACACA 410
DB	7612 AATTTATTTAGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7671
QY	411 TTTTATTTTAAAGTGTGCAATAATTTTCTTAATTTAGTTTAAATTTCTAATCTTTTACTTTT 468
DB	7672 TTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7731
QY	469 CTCTCATCTAAATTTAATCTTATTTACCATAATTTAAATTTGTCGAGACAAATTTATTTT 528
DB	7732 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7791
QY	529 AATCTCACCTCCATTAATGATATTTATTAATTTTGTTCGA--TACTTCTTATTTTCACT 586
DB	7792 TATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7851

RESULT 8
ADB54318
ID ADB54318 standard; DNA; 11222 BP.
XX AC ADB54318;
XX DT 04-DEC-2003 (first entry)
XX DE Pretreated genomic DNA region 242.
XX KW colon cell proliferative disorder; non methylated CpG dinucleotide;
XX KW cytosine; cancer; adenoma; carcinoma; cytosine methylation state; ds.
XX OS Unidentified.
XX SN WO2003072821-A2.
XX PD 04-SEP-2003.
XX PF 27-FEB-2003; 2003WO-EP002035.
XX PR 27-FEB-2002; 2002EP-00004551.

ABL34155	ID	ABL34155 standard; DNA; 15548 BP.			
XX	AC	ABL34155;			
XX	DT	26-MAR-2002 (first entry)			
XX	DE	Human immune system associated gene SEQ ID NO: 2128.			
XX	KW	Human; immune system disease; cytosine methylation; antiasthmatic; antiarteriosclerotic; antianaemic; cytostatic; nootropic; neuroprotective; anti-HIV; anticonvulsant; ophthalmological; anticrimeumatic; antiarthritic; antididiabetic; antipsoriatic; antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia; acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; ds.			
XX	OS	Homo sapiens.			
XX	PN	WO200200928-A2.			
XX	PD	03-JAN-2002..			
XX	PF	02-JUL-2001; 2001WO-EP007537.			
XX	PR	30-JUN-2000; 2000DE-01032529.			
XX	PR	01-SEP-2000; 2000DE-01043826.			
XX	PA	(EPIG-) EPIGENOMICS AG.			
XX	PI	Olek A, Piepenbrock C, Berlin K;			
XX	DR	WPI; 2002-130909/17.			
XX	NK	Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation.			
PS	Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.				
XX	The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory ulcerative bowel diseases. The present sequence is a gene of the invention				
SQ	Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 U; 0 Other;				
Query Match	9.8%; Score 96.2; DB 6; Length 15548;				
Best Local Similarity	53.7%; Pred. No. 0.00014;				
Matches	267; Conservative 0; Mismatches 223; Indels 7; Gaps 3;				
QY	90	TATTGCATTTACTCTCATTTCATTATTATTAAATTATTATAATTGGTGAATAAATT 149			
Db	11480	TAGTAATTTTATATATATTTTTTTTATATTTTATTTTATTTATTTTATTTTATTT 11539			
QY	150	CSTAATACTTTACTTT-TTTTTAAAAAGAATTTATTATAATTTTATATTTTAGATAA 208			
Db	11540	TTATTATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 11599			
QY	209	AAATCTAATACCTTACTTTTTTTTTTAAAAGAATTCGAATCGGTCTTTTCTTAATTA 268			
Db	11600	ATATTTTATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 11659			
QY	269	GTTTAAATCTACTACTAATATAAAATTCGATCGGATTAGTGCGTCAAGTCAAG 328			
Db	11660	TATTTATTTTATTTTATTTTATTTT---ATTATTTATTTTATTTTATTTTATTT 11716			
QY	329	TCACATGAATTTTGTGGAGAAAAATAAAATTAACAACATTTTTCGATTAAATTTATTA 388			

[illegible]

Search completed: March 14, 2005, 23:50:30
Job time : 575 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 20:13:50 ; Search time 4297 Seconds
(without alignments)
11107.374 Million cell updates/sec

Title: US-10-089-543-2_COPY_449_1433

Perfect score: 985
Sequence: 1 gatgtgaccagtggttgag.....tgagagaatccttcacatc 985

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hhg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_ste:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	985	100.0	1433	6	AX370650	Sequence
2	981.8	99.7	5765	8	AF487511	Gossypium
3	119.2	12.1	170627	2	AC125567	Rattus no
C 4	113.2	11.5	86826	3	PFWAL3P5	AL034556 Plasmodiu
C 5	112.8	11.5	67970	3	PFWAL1P3	AL031746 Plasmodiu
C 6	112.2	11.4	180861	2	CR391970	CR391970 Danio rer
7	111.2	11.3	286208	2	AC117140	AC117140 Rattus no
C 8	110.2	11.2	258658	3	AE014832	AE014832 Plasmodiu
C 9	110	11.2	191191	2	CR847965	CR847965 Danio rer
C 10	109.8	11.1	348174	3	CR382399	CR382399 Plasmodiu
C 11	109.6	11.1	169056	2	CR536620	Danio rer
C 12	109.6	11.1	213093	5	BX247875	BX247875 Zebraphis
13	109.2	11.1	110000	2	PFWAL8P1_12	Continuation (13 o
C 14	109.2	11.1	179951	2	CR392343	CR392343 Danio rer
C 15	109.2	11.1	246611	2	AC111404	AC111404 Rattus no
C 16	108.8	11.0	145925	5	BX842571	BX842571 Zebraphis
C 17	108.4	11.0	110000	2	BX927162_2	Continuation (3 of
C 18	108.4	11.0	192182	5	BX649602	BX649602 Zebraphis
19	108.2	11.0	140289	5	BX005458	BX005458 Zebraphis

20	107.8	10.9	584	3	AF201315	AF201315 Dugesia p
C 21	107.8	10.9	145962	2	CR352215	CR352215 Danio rer
22	107.8	10.9	156404	2	CR626908	CR626908 Danio rer
C 23	107.6	10.9	27694	8	HASMT	D31785 Pichia cana
24	107.4	10.9	8056	6	AX599046	AX599046 Sequence
C 25	107.2	10.9	138342	2	CR376794	CR376794 Danio rer
26	106.6	10.8	165797	2	CR847851	CR847851 Danio rer
27	106	10.8	125632	2	CR753868	CR753868 Danio rer
28	106	10.8	186431	2	AC022281	AC022281 Homo sapi
29	105.8	10.7	8056	6	AX598900	AX598900 Sequence
30	105.8	10.7	166719	9	AC136957	AC136957 Homo sapi
31	105.8	10.7	198434	9	AC109994	AC109994 Homo sapi
32	105.8	10.7	250029	3	AE014839	AE014839 Plasmodiu
C 33	105.2	10.7	83435	5	BX640469	BX640469 Zebraphis
C 34	105	10.7	131682	9	AL672277	AL672277 Human DNA
C 35	104.8	10.6	162445	9	AL158151	AL158151 Human DNA
36	104.6	10.6	14867	3	AE001398	AE001398 Plasmodiu
37	104.6	10.6	172777	2	CR847809	CR847809 Danio rer
38	104.6	10.6	196784	9	AC016543	AC016543 Homo sapi
C 39	104.6	10.6	261771	2	CR356223	CR356223 Danio rer
40	104.4	10.6	4865	3	AF364131	AF364131 Anopheles
C 41	104.2	10.6	153751	3	AC116551	AC116551 Dictyoste
42	104.2	10.6	169357	2	CR847871	CR847871 Danio rer
C 43	104.2	10.6	209719	2	CR847825	CR847825 Danio rer
44	104.2	10.6	227211	2	CR847561	CR847561 Danio rer
C 45	104	10.6	104992	2	AC005504	AC005504 Plasmodiu

ALIGNMENTS

RESULT 1
AX370650
LOCUS AX370650 1433 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 2 from Patent WO0210377.
ACCESSION AX370650
VERSION AX370650.1 GI:19168815
KEYWORDS
SOURCE Arabidopsis sp.
ORGANISM Arabidopsis sp.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Cai, L.I., Li, X., Cheng, N. and Liu, J.W.
TITLE Isolation and characterization of a fiber-specific_g(b)-tubulin promoter from cotton
JOURNAL Patent: WO 0210377-A 2 07-FEB-2002;
Institute of Molecular Agrobiology (SG)
FEATURES
source
1..1433
/organism="Arabidopsis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:29726"

ORIGIN	Query Match	Best Local Similarity	Score 985;	DB 6;	Length 1433;
	Matches 985;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	GATGTGACGAGTGGTGGAGAGAGAGGCTACCGATTGGTCAAGTGGGACCAATTTTAA 60				
Db 449	GATGTGACGAGTGGTGGAGAGAGGCTACCGATTGGTCAAGTGGGACCAATTTTAA 508				
QY 61	TTTTACCTCCTCGCTAGATTTCGTAATACTATTGCTATCTCATTTTCATTATTATT 120				
Db 509	TTTTACCTCCTCGCTAGATTTCGTAATACTATTGCTATCTCATTTTCATTATTATT 568				
QY 121	AATTATTTTATATTATTTGGATAAAAAATCTAATACTTTTACTTTTAAAAAAGAAATTT 180				
Db 569	AATTATTTTATATTATTTGGATAAAAAATCTAATACTTTTACTTTTAAAAAAGAAATTT 528				
QY 181	ATTTAATTTATTTTATTTATTTAGATAAAAAATCTAATACTTTTACTTTTAAAAAAG 240				

Db 629 ATTAAATATTATATATATATAGATAAAATCTCAATCTTACTTTCTTTTAAAAAG 688
QY 241 AATTCAATGGGTTTTCTTAATAGTTAAATCTAFACTAATTAATAAAATCTTG 300
Db 689 AATTTCAATGGGTTTTCTTAATAGTTAAATCTAFACTAATTAATAAAATCTTG 748
QY 301 ATCGGATAGTGGTGCTCAAGTCAAGTCAATGATTTCTGAGAGAAAAATAAAA 360
Db 749 ATCGGATAGTGGTGCTCAAGTCAAGTCAATGATTTCTGAGAGAAAAATAAAA 808
QY 361 TTAACAACATTTTTCGATTAATTAATATATATAATAATAATAACACATTTTATTTA 420
Db 809 TTAACAACATTTTTCGATTAATTAATATATATAATAATAATAACACATTTTATTTA 868
QY 421 ATGTTGTCATATATATTTTAAATTAATTTTCAAGCAACAAATTTACATCTCATCATTA 480
Db 869 ATGTTGTCATATATATTTTAAATTAATTTTCAAGCAACAAATTTACATCTCATCATTA 928
QY 481 AATTTAATCTTATACCAATTTAAATTTGTGAGCAATTTATTTTAAATCTCACCTTC 540
Db 929 AATTTAATCTTATACCAATTTAAATTTGTGAGCAATTTATTTTAAATCTCACCTTC 988
QY 541 CATTAATGCAATATTAATTTTGTGATATCTTATTTTCACTCTCTTAAACATTAATCA 600
Db 989 CATTAATGCAATATTAATTTTGTGATATCTTATTTTCACTCTCTTAAACATTAATCA 1048
QY 601 TTAACCAATTTTGAATGTTATATTTTAAATTTTCACTCTTATTTTGTGCTCTGGTCC 660
Db 1049 TTAACCAATTTTGAATGTTATATTTTCTTAACTTATTTTCACTCTTGTGCTCTGGTCC 1108
QY 661 ATCTGGAAGGCCACCGTCCAGCTCTCAACACACACTTTGGCACATCAATTTCCAGT 720
Db 1109 ATCTGGAAGGCCACCGTCCAGCTCTCAACACACACTTTGGCACATCAATTTCCAGT 1168
QY 721 AACTACATGTTTACGTTACTAAGCAAAATCCCAATTTTCAAAATTTCCAGGAAA 780
Db 1169 AACTACATGTTTACGTTACTAAGCAAAATCCCAATTTTCAAAATTTCCAGGAAA 1228
QY 781 ACGAAGCGTCCGTTACTAAGCAAAATCCCAATTTTCAAAATTTCCAGGAAA 840
Db 1229 ACGAAGCGTCCGTTACTAAGCAAAATCCCAATTTTCAAAATTTCCAGGAAA 1288
QY 841 ATCTTTAACTCTCTATATAACCCAAACCACTCTCATCACCATTTCCCATATAAAGA 900
Db 1289 ATCTTTAACTCTCTATATAACCCAAACCACTCTCATCACCATTTCCCATATAAAGA 1348
QY 901 ATTTCCGGAATCTTATCTTTTATATTTTCTCTCCAAATTTCCGTCACCTTCGGA 960
Db 1349 ATTTCCGGAATCTTATCTTTTATATTTTCTCTCCAAATTTCCGTCACCTTCGGA 1408
QY 961 GAAATGAGAGAAATCTTCCATC 985
Db 1409 GAAATGAGAGAAATCTTCCATC 1433

RESULT 2
LOCUS AF487511 5765 bp DNA linear PLN 17-OCT-2002
DEFINITION Gossypium hirsutum beta-tubulin (TUB1) gene, complete cds.
ACCESSION AF487511
VERSION AF487511.1 GI:19569608
KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM
Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 5765)
Li, X.B., Cai, L., Cheng, N.H. and Liu, J.W.
Molecular Characterization of the Cotton GhtUB1 Gene That Is
Preferentially Expressed in Fiber
Plant Physiol. 130 (2), 666-674 (2002)

MEDLINE 22263995
PUBMED 12376634
REFERENCE 2 (bases 1 to 5765)
AUTHORS Li, X.B., Cai, L. and Liu, J.W.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2002) Laboratory of Plant Reproduction and Cotton
Biotechnology, Institute of Molecular Agrobiolgy, 1 Research Link,
NUS, Singapore 117604, Republic of Singapore
FEATURES
source 1. 5765
/organism="Gossypium hirsutum"
/mol_type="genomic DNA"
/db_xref="taxon:3635"
<1377..>3349
/gene="TUB1"
join(<1377..1771,1876..2144,2676..>3349)
/gene="TUB1"
product="beta-tubulin"
join(1377..1771,1876..2144,2676..3349)
/gene="TUB1"
note="microtubule protein subunit"
/codon_start=1
/product="beta-tubulin"
/db_xref="GI:19569608"
/translation="MREILHIQGGCGNQIGAKFEWVVCABHGIDSTGRYGGDELQ
ERINYYNEASCGRFVPRVLMLEPGTMDSRGPGYQIPRPNFVFGGAGNNMA
KGHTGEALIDSVDVVRKEAENCDCLOGFQVCHSGGGTSGMGLLILKIRBEYP
DRMLTSPVPSPKVSDTVPEYNATLSHOLVENADECMVLNELYDLCFRFLKLT
TPSPGDLNHLISATMSGVTCCLRPGOLNSDLRLKLVNLI PPRLHFMVGFALTSR
GSQYRALTVFELTQMDAKNMCAADPHRGVLTASAVFRGKMSKEVDEQMINVO
KNKSSYFVEWIPNNKSTVDIPPIGLKMASTFIGNSTSIQEMFERRVSEQFTAMPRK
AFLHWYTBGMDMEFTEAESNMNDLVSEYQQYQDADDEYEEEEEYEA"

ORIGIN
Query, Match 99.7%; Score 981.8; DB 8; Length 5765;
Best Local Similarity 99.8%; Pred. No. 2.7e-117;
Matches 983; Conservative -0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GATGTGACCAAGTGGTTGGAGAGAGAGGCTCTACCGATTGGTCAAGTGGCAACCAATTTT 60
Db 413 GATGTGACCAAGTGGTTGGAGAGAGAGGCTCTACCGATTGGTCAAGTGGCAACCAATTTT 472
QY 61 TTTTACCTCTGCTCTAGATTCGTAATACTATTCGATTTATCTCATTTTATTTATTT 120
Db 473 TTTTACCTCTGCTCTAGATTCGTAATACTATTCGATTTATCTCATTTTATTTATTT 532
QY 121 AATTAATTTTATATTTTGGATATAAAATTTCTAATACTTTTCTTTTAAAAAGAAATTT 180
Db 533 AATTAATTTTATATTTTGGATATAAAATTTCTAATACTTTTCTTTTAAAAAGAAATTT 592
QY 181 AATTAATTTTATATTTTAGATAAAATTTCTAATACTTTTCTTTTAAAAAG 240
Db 593 AATTAATTTTATATTTTAGATAAAATTTCTAATACTTTTCTTTTAAAAAG 652
QY 241 AATTTCAATGGGTTTTCTTAATAGTTAAATCTAFACTAATTAATAAAATTTCTG 300
Db 653 AATTTCAATGGGTTTTCTTAATAGTTAAATCTAFACTAATTAATAAAATTTCTG 712
QY 301 ATCGGATAGTGGTGCTCAAGTCAAGTCAATGATTTCTGAGAGAAAAATAAAA 360
Db 713 ATCGGATAGTGGTGCTCAAGTCAAGTCAATGATTTCTGAGAGAAAAATAAAA 772
QY 361 TTAACAACATTTTTCGATTAATTAATATATAATAATAATAACACATTTTATTTA 420
Db 773 TTAACAACATTTTTCGATTAATTAATATATAATAATAATAACACATTTTATTTA 832
QY 421 ATGTTGTCATATATTTTAAATTAATTTTCAAGCAACAAATTTACATCTCATCATTA 480
Db 833 ATGTTGTCATATATTTTAAATTAATTTTCAAGCAACAAATTTACATCTCATCATTA 892
QY 481 AATTTAATCTTATACCAATTTAAATTTGTGAGCAATTTATTTTAAATCTCACCTTC 540


```
Db      893  AATTTAATCTTATTACCATTAATAAATGTCGAGGACAATTAATTTTAAATCTCACCCCTC 952
Qy      541  CATTAAATGCATATTATTAATTTTGTGTCATACCTCTTATTTTCACTCTCTAAACATTAATCA 600
Db      953  CATTAATGCATATTATTAATTTTGTGTCATACCTCTTATTTTCACTCTCTAAACATTAATCA 1012
Qy      601  TTAACCCAAATTTGAACCTGTTAATTTTCTTAACCTTATTTCACTATTTGGCTCTGGTCTCC 660
Db      1013  TTAACCCAAATTTGAACCTGTTAATTTTCTTAACCTTATTTCACTATTTGGCTCTGGTCTCC 1072
Qy      661  ATCTGGAAGGCCACCGCTGTCACCAACGACATTTTGCACGTCATCAATTTCCAGT 720
Db      1073  ATCTGGAAGGCCACCGCTGTCACCAACGACATTTTGCACGTCATCAATTTCCAGT 1132
Qy      721  AACTACATTTGTACAGTTACTAAGCAAAATCCCAATTTTCAAAAATTTCAATTTCCCGAGGAA 780
Db      1133  AACTACATTTGTACAGTTACTAAGCAAAATCCCAATTTTCAAAAATTTCAATTTCCCGAGGAA 1192
Qy      781  ACGAAACGTCGTTACTAAGCAACCTAAACCGAGCTCAACCTGCGGTCAATTAACGGAA 840
Db      1193  ACGAAACGTCGTTACTAAGCAACCTAAACCGAGCTCAACCTGCGGTCAATTAACGGAA 1252
Qy      841  ATCTTTTAACCTCTCTATATAACCAACCAACCTCTCTATCAACCATTTCCCAATAAAGA 900
Db      1253  ATCTTTTAACCTCTCTATATAACCAACCAACCTCTCTATCAACCATTTCCCAATAAAGA 1312
Qy      901  ATTTCGGGAATTTCTTATCTCTTTATATTTTCTCTCCAAATTTCCCGTCATTTCCCGA 960
Db      1313  ATTTCGGGAATTTCTTATCTCTTTATATTTTCTCTCCAAATTTCCCGTCATTTCCCGA 1372
Qy      961  GAAATGAGAGAAATCTTTCATC 985
Db      1373  GAAATGAGAGAAATCTTTCATC 1397

RESULT 3
AC125567
LOCUS
DEFINITION
Rattus norvegicus clone CH230-9H22, *** SEQUENCING IN PROGRESS ***,
4 unordered pieces.
ACCESSION
AC125567
VERSION
AC125567.5 GI:24817949
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 170627)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,J., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Jackson,L., Jacobs,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
```

```
Lorensuheva,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,M., Martinez,E.,
Mayhew,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaekelimeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plapper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,M., Sodergren,E., Song,X.-Z., Sorelle,R., Soosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlecsyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 170627)
Worley,K.C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 170627)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23096537.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with 'Ns' to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDDV
Center clone name: CH230-9H22
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 162619 bases at least Q40
Consensus quality: 164392 bases at least Q30
Consensus quality: 165269 bases at least Q20
Estimated insert size: 169231; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
```

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* this record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 109158: contig of 109158 bp in length
* 109159 109258: gap of unknown length
* 109259 110280: contig of 1022 bp in length
* 110281 110380: gap of unknown length
* 110381 111540: contig of 1160 bp in length
* 111541 111640: gap of unknown length
* 111641 170627: contig of 58987 bp in length.

FEATURES

source

1..170627
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9H22"

misc_feature

1..1284
/note="wgs end extension
clone_end:Sp6"

misc_feature

/note="clone boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH341988"

ORIGIN

Query Match 12.1%; Score 119.2; DB 2; Length 170627;
Best Local Similarity 50.2%; Pred. No. 6.4e-07;
Matches 295; Conservative 0; Mismatches 293; Indels 0; Gaps 0;

QY 58 TTATTTTACTCTGCTAGATCGTAATACTATGCACTTTATCTCATTTTCATTTATTTA 117
DB 28563 TTTTTTTTATTTTTTATATTAATATTTTTTTTATTAATATTTTTTTTATTTATTTT 28622
QY 118 TTTAATTTATTTATTTATTTTGGATAAAAATCTAATCTTTACTTTTTTTTAAAAAGAA 177
DB 28623 TTTTTTTTATTTTTTTTTTTTAAAAATTTTTTATATTTTTTTTTTTTTTTTTTTTT 28682
QY 178 TTTATTTAATTTATTTATTTATTTAGATAAAAAATCTAATCTTTACTTTTTTTTAA 237
DB 28683 TTTAATTTTAAATATTTATTTTTTTTAAATTTTTTTTTTTTATTTTTTTTTTATAT 28742
QY 238 AAGAAATTCATTTGCGTTTTTTCTTAATTTAGTTTTAATCTATATACTAATATAAAAT 297
DB 28743 TTTTATTTTTTTTTTTTTTTTATATATATATATTTTTTTTATTTTATTAATATTTT 28802
QY 298 CTGATCGATTAGTGTGCTCAAGTCACATGAATTTGTTGGAGAAAAATAA 357
DB 28803 TTTTATAATTAATAATTTTTTTTATTTTATTTTAAATAATATTTTAAATAATTTATTA 28862
QY 358 AATTTAAACACATTTTTCGATTAATTTATATATATATATATATATAACACATTTTAT 417
DB 28863 TTATTATATAAAATATTTTTTTTTTTTTTTTTTTTTTTTATTTATTTATTTTAAATAT 28922
QY 418 TTTAATGTGTCAATAATTTTTTTAATTAATAATTTTTCAGCACAAATTCACATCTCATCA 477
DB 28923 TTTTATTTATATTTTTTTTTTTTATTTATTTATTTTATTTTAAATATATATTTATTT 28982
QY 478 TTTAAATTTATCTTATTACATAATTTAAATTTGTGAGGACAAATTTTTTTTAACTCACC 537
DB 28983 TATTAATTTTTTTTTTATATATATATATATTTTTTTTTTAAATTTTTTTTATTTTAT 29042
QY 538 CTCATTAATGCATATATTTAATTTTTTGTGCGACTCTTATTTTACCTCCCTAACATTA 597
DB 29043 TTTAAATTTATTTATTTATTTTATTTTTTTTATTTTATTTTATTTTATTTATTTAT 29102
QY 598 TCATTAACCAATTTTGAACCTGTATAATTTCTTAACCTTATTTTACCTAT 645
DB 29103 TTTTTCATTTTTTTTTTTTTTTTTTAAATATTTTATTTATTTATTTATTTT 29150

RESULT 4

PFMAL3P5/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

PFMAL3P5 86826 bp DNA linear INV 29-JAN-2003
Plasmodium falciparum MAL3P5, complete sequence.
AL034556 AL008971 AL008972 AL008978 AL010141 AL010153 AL010162
AL010206 AL010210 AL139179 AL844502
GI:23477013
HTG; centromere; CTRP protein; initiation factor E4;
Serine/threonine protein phosphatase.
Plasmodium falciparum 3D7
Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 86826)
Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T.,
Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T.,
Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S.,
Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J.,
Moule, S., Mungall, K., Murphy, L., Oliver, K., Price, C., Quail, M.A.,
Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S.,
Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and
Barrell, B.G.

The complete nucleotide sequence of chromosome 3 of Plasmodium

falciparum

Nature 400 (6744), 532-538 (1999)

99376085

10448855

2

Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,
Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,
Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Corton, C.,
Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,
Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,
Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,
Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,
Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D.,
Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N.,
Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,
Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A.,
Rabbinowitz, E., Rajandream, M.A., Rutter, S., Rutherford, K.M.,
Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,
Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,
Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and
Barrell, B.G.

Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13

Nature 419 (6906), 527-531 (2002)

22255708

12368867

3 (bases 1 to 86826)

Lawson, D., Bowman, S. and Barrell, B.

Direct Submission

Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge

UK

On Oct 2, 2002 this sequence version replaced gi:7711064.

For more information about this sequence or the Malaria Project,

see http://www.sanger.ac.uk/Projects/P_falciparum.

Location/Qualifiers

1..86826

/organism="Plasmodium falciparum 3D7"

/mol_type="genomic DNA"

/isolate="3D7"

/db_xref="taxon:36329"

/chromosome="3"

/clone="MAL3P5"

join(322..603,826..1023,1197..1301,1458..1533,1696..2942)

/gene="MAL3P5.1"

/note="synonym: PFC0575w"

join(322..603,826..1023,1197..1301,1458..1533,1696..2942)

/gene="MAL3P5.1"

/note="Possible signal sequence present, revised; added
new exon 2, revised; incorporated new intron in last exon
in comparison with P. knowlesi and P. yoelii"

/codon_start=1

```
/product="hypothetical protein"
/protein_id="CAB38969.3"
/db_xref="GI:23477014"
/db_xref="UniProt/TREMBL:O97258"
/translation="MYLKNVYIVISSCFILFDLCRSFHLKKYKXNMNMKSVTFPL
RSPQYRKFQKSRKNVSKKQKPLFENLKKGFSLGFRNQYQKVIDDVIS
NINLRIQVTHKKSNSEFTENIKQILLHCVFSLDKFINNLSIILKHFQMSNIT
VSHLQISEKVEKKAENYALHILFLKDENTILFSMMHMDFFSKQKVICIRD
IKRYIECVGLDIRKIRYLNCEKKIKFMDKDAIKMWEINMKDDLYFNHYHDE
LRCFTMKNIERNKNIIRSNVDNINDDISDKMDAKMNPIDVINNI SLDEKIKEQ
HNPDDNKKDDYEQQLFNDNIIVKIEEDQPLYNLNNNNNNNTWKNK
FKNTDYNLDDYVEKEDLVIOKNIDDIYIKNTIGMKSLEEFKNQIFEQADLEPQ
NELSNVNLQHGKVSNDENTKSTEHKKNVINKGYDTLQNMENNFIKGINDN
ISNDNNEINNIILKKLQNSDEINLTSDIYERLTKVLTIOKIEYLKPYQYDI
INEQYPIRKEKTYDILLNGYKIVMSPDVNSLPEKTKIDISIPNEKKNQMNQKN
SKNYIIMFNMV"
complement (3354..6644)
/gene="MAL3P5.2"
/notes="synonym: PFC0580c"
complement (3354..6644)
/gene="MAL3P5.2"
/notes="Signal peptide predicted by SignalP 2.0 HMM (signal
peptide probability 0.635, signal anchor probability 0.287)
with cleavage site probability 0.594 between residues 21
and 22; CDS conserved in P. knowlesi and P. yoelii"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB38968.1"
/db_xref="GI:4493932"
/db_xref="UniProt/TREMBL:O97257"
/translation="MUGLRKNVYLLVSPSLFAYFLKRGKDNENNYETLINNDIE
KIKRIHNNKSYIPLLFNIYDSYIKNKLRLWYLFKPKRRKQKEEYIYITNVRK
KREAIKNFIQDEQLFNKFIYEIVNEHTQAPNTLPHNNNNKPKPLDNIHSC
SLYYNKLNNKLNELKELFYKQINNEHTQAPNTLPHNNNNKPKPLDNIHSC
KNTNISSTYNNMEKENIYDKYNIHNFYTESISYKDCQCHITLWNIYLNQTY
DNIICRLTNTYNIYFNMILKVIYKWEIILNLYNHIEDMKKIKQKNNTYS
LPKISYFFFKKEENIYDLPEQNMHLKENDCKFYNSNENTHNNIYKISDNY
FYDHNNSNRCSFKNLKQDQDNTKTHIMGEKYPMNKSDEKKNNTCNINIEK
DQKODILKIKYELKGNKLDIOIQLNELYMIYMLLEFCSLKILIKKNIHLKEME
PQDKNIIVLNGADYMNLRILKFSKNEERENINSPASFFLSKNIIVFEDEIG
RSDNTIYNNVDKENTTKTTNNNDNDNICSNNDHICSNNDHICSNNDHICSN
NNNICSNNNNNICSNNNNICSNNNNICSNKMLDEFQCKNKFNDYTRKKEKRI
YELAKIYNNIYDLKKEKKEHONEDNTINLIYIKKFPWIFLYKNIINKDTSFIEH
NNIYVNGDIKNNNIIPKKYNEFESSIISFYKDIYENYKRLIYIDNIIRKFCR
YFLKWEHNLRLKMKRAFYIYNFDOFIINNYHIHKKNIHKKIHLKQCKDE
IDIVKPDLYYCMINNIINNI FSVIHKVDHNECYRIFKAYNKI LLYEYNLMEKENIY
YKNIKKVLYTLNNISNDLYPVNIYKNIYKQKYNKPNFSHIFYSKNDIHLILF
LYTORIGNCDDISYIYSKYNFNKPNFLNLYIYELHYIYSEKKKKKKFFSISSP
YSYDQVNSFTSFSYFSSLYLLFLILFHPMDYASYIIPKTIYISGLPTYYSLYNNI
MVYCGPKT"
join (7669..7717,7800..7829,7912..7940,8064..8106,
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
9369..9505,9613..9838)
/gene="PFC0581w"
join (7669..7717,7800..7829,7912..7940,8064..8106,
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
9369..9505,9613..9838)
/gene="PFC0581w"
/notes="Revised: new gene prediction, splicing prediction
very tentative"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB90285.1"
/db_xref="GI:7711066"
/db_xref="UniProt/TREMBL:Q9NPA4"
/translation="MKKKGHYFSIKFVNDKKNVLSKQYKIRIINWIMGLHIFRL
QKIFHYFAMQIASFFSHKRYRKVEDIIDGIEKWFNTIDYLSIESIQDNDFS
NDIKTTVYAKMKQQLTITSYNNKIDVTNASFQWASQPEYIFLNIFKFSHRWSPGA
LKVQDEEIKWNNKLSFSALSDNSVTKKIYVDLTNLIIESETKYNFASGVGVV
TLKKEKIKTNLKSKEYKPNQVWDMKKEKRIHLITPTVINLFFLSLSHYHDSV
QNFLKEKNSKLQDDIDEDEKDFDEILREAKKSEYEDKDEEL"
complement (join (10024..10044,10320..10336,10397..10445,
10613..10674,10773..10798,10929..10999,11157..11237,
```

```
11453..11515,11715..11767,11950..12040))
/gene="PFC0582c"
complement (join (10024..10044,10320..10336,10397..10445,
10613..10674,10773..10798,10929..10999,11157..11237,
11453..11515,11715..11767,11950..12040))
/gene="PFC0582c"
/notes="Revised: new gene prediction"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB90286.1"
/db_xref="GI:7711067"
/db_xref="UniProt/TREMBL:Q9NPA3"
/translation="MTTSEYFDDYEDLNRHLKHKIYINTKEDNKNKREYIPE
ASTCKYQINIEINSLPKSNKIINQINTVLDLKKYNTVQKSDADYSEBYLNSQY
DMTKYKEGVDFLKESERRAQIEDIGYTIMSELNSQRSAILRTHKHTDETROENRV
KMLLIITIANIGNSFFLF"
complement (join (13653..17754,18005..19353))
/gene="MAL3P5.4"
/notes="synonym: PFC0590c"
complement (join (13653..17754,18005..19353))
/gene="MAL3P5.4"
/notes="Signal peptide predicted by SignalP 2.0 HMM (signal
peptide probability 0.921, signal anchor probability 0.075)
with cleavage site probability 0.888 between residues 20
and 2, revised: changed donor in comparison with P.
knowlesi"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB38986.3"
/db_xref="GI:23477015"
/db_xref="UniProt/TREMBL:O97275"
/translation="MLFYKLCILPLSFIIIIHGHKDISSTKLLKSGNLNNKIKKGR
RYGKYNSKTYIEKKRYNIKKQNNNSYFCNVYKNDVNNNTAYNYINNP
KNLKEYEKIKNHYIKKKKIFSLKFSQNRKNEKKYFFNFTHDIDKNIKVL
NIDYLENNIYYSIKSEKTPITKIVLLGTELLSVLJHMKNVYKLILDFDKLIFK
GEIWLFPYIYIGNLYQYILMPNLYNIMSSVEISHYKPEDFLIFLFGYSNLL
FTIWMANNENIMNVKLIYHNPKNFFIKDCVSKTSRSSTNNNNNNNNNNSSNNL
HYNNSKNIDKKEQYHLYVESTYLYWSRINEGTILNCFELFIEKAYVPPFFII
QNIILNFSLYFVASIFSSYLYFTEKYFKENYLRFLFKLLKVIHLYPLVDRYKE
KPNFLQKQYIHNVPRIHTMKNWISDLKNIENCKNKKVKNKSNFLVLIILK
LPKEKIVQDPMIILKLSKYIKIEKVLVYICQIHEIYKFPTRDTLLIILK
KNFDNIYINLISKILMKNNKMSYKOLALIYLSKNIYLTDEQYKKEIPNFSIL
KPEHNNVNLNLSLSLFFSYSVYFNNCFYIYFHSFNNITKFINIKNNLHL
YNSTDLNLYTGLSHNNYTPNDQNKQNNNNNNNNNNNNNNNNNNNNNNNN
DHLRSDNFIQNDYDQGEYKHIQONKHIQONKHIQONKHIQONKHIQONKHI
NSNNQQLIQONDEKENKLIYDNSKNNLLKDCOHNIVEKDGKKONILKNIINIK
KIEKLSFKIQLIYVNLVSLNKNILNKYKYPHQAQKINIRYINIVYNKIY
LNDEBNTSCHILKIKNDKNDKILYHDMKFLYNLNNIYIRNDLNNMKQIILLYGL
KPNPMFLOFEKILIKRFICLPKSIQIGKEEIMFLYQYFFVTRCLFNLKKQNNLF
ISQDEYENIYISDKYNESAKLDSNNPNNLKERNNHGGKNDTLDYIHDLPYK
NKKKRDYKLYDNFIYFPAYYVEQKDHIDYNESVNNFNNKSFOLKCKKLNK
NNNNNNNNNNIYIDTNIOTVKNYSCTHNNVKNETNDYNNSTIRNOHPNDQV
ILANNPFFYNNKLVVSDIDEVELTCYNVLDIYKIVCLKLLTLKNNKLSCLQSID
ILCIYEKAMIRDYRIIKLYNKKELLYDNTLYLKVNIINIVKPNLYNMISYIQINKI
LTFINNNINISIQILKGLMISVHKHKNKJSPFHNNLNQVNAANYLPKNLYNQNI
QDLKIEWNVYDNLTFKFLKNIILNIVKRYVQNCNYSKNYKNTHTNNLNKNEQ
```


Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1. (bases 1 to 180861)
Sims, S.
Direct Submission
Submitted (25-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 27, 2004 this sequence version replaced gi:46559024.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk
----- Project Information
Center project name: ZK5FP22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 177004 bases at least Q40
Consensus quality: 177517 bases at least Q30
Consensus quality: 178043 bases at least Q20
Insert size: 180161; sum-of-contigs
Insert size: 176592; 4.4% error; agarose-fp
Quality coverage: 11.58x in Q20 bases; sum-of-contigs Quality
coverage: 11.81x in Q20 bases; agarose-fp

REFERENCE AUTHORS TITLE JOURNAL

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1
* 4193 4192: contig of 4192 bp in length
* 4293 4292: gap of 100 bp
* 15167 15166: contig of 10874 bp in length
* 15167 15266: gap of 100 bp
* 15267 33335: contig of 18069 bp in length
* 33336 33435: gap of 100 bp
* 33436 56531: contig of 23096 bp in length
* 56532 56631: gap of 100 bp
* 56632 79397: contig of 22765 bp in length
* 79397 79496: gap of 100 bp
* 79497 82394: contig of 2898 bp in length
* 82395 82494: gap of 100 bp
* 82495 101952: contig of 19458 bp in length
* 101953 102052: gap of 100 bp
* 102053 180861: contig of 78809 bp in length.

FEATURES source

1..180861
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-95F22"
/clone_lib="DanioKey"
1..4192
/note="assembly fragment:00026
fragment chain:1"
4293..15166
/note="assembly fragment:00087
fragment chain:1"
15267..33335
/note="assembly fragment:00301
fragment chain:1"
33436..56531
/note="assembly fragment:00964
fragment chain:1"
56632..79396
/note="assembly fragment:00631
fragment chain:1"
79497..82394

misc_feature /note="assembly_fragment:00011"
82495..101952
misc_feature /note="assembly_fragment:01300.0"
102053..180861
/note="assembly_fragment:01684"

ORIGIN

Query Match 11.4%; Score 112.2; DB 2; Length 180861;
Best Local Similarity 54.0%; Pred. No. 5e-06;
Matches 318; Conservative 0; Mismatches 263; Indels 8; Gaps 4;
QY 52 AATTTTATTTTACCTCCTCCCTAGATTCCTAAATCTACTATTCATTTATCTCATTTTCAT 111
DB 82337 ATTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 82278
QY 112 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 171
DB 82277 TAAAT-TTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 82219
QY 172 AAAGAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 231
DB 82218 TAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 82160
QY 232 TTTAAAGAAATTTCAATTCGTTTTCCTTAAATTTAGTTTCTTAAATTTCTTAAATTTCT 289
DB 82159 AATAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 82100
QY 290 TAAAAATTTCTGATCGGATTAGTGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 349
DB 82099 TATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 82040
QY 350 AAAAAATAAAAAATTAACACATTTTTCGATTAATTTTATTTATTTATTTATTTATTTATTTA 409
DB 82039 AAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 81984
QY 410 ATTTTATTTTAAATGTTGTCAATAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTT 469
DB 81983 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 81924
QY 470 TCTCATCAATTAATTAATCTTATTTACCAATAATTAATAATTTGAGGACAAATTTTATTT 529
DB 81923 TTAATATAATAATTTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 81864
QY 530 ATCTCACCTCCATTAATGCAATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTAT 589
DB 81863 ATATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 81804
QY 590 ACATTAATCAATTAACCAATTTTTCGAATCTGTTATTTATTTTATTTTATTTTATTTAT 638
DB 81803 TTTTATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 81755

RESULT 7

AC117140 286208 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-36587, WORKING DRAFT SEQUENCE, 9
unordered pieces.
AC117140
AC117140.4 GI:24635509
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 286208)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burrell, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Lousegod, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Maindardine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwaokeme, O., Okwunnu, G., Olarnpunsagoon, A., Pal, S., Pank, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Steed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Stelmle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valdes, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 286208)
Worley, K. C.

Direct Submission
Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 286208)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 6, 2002 this sequence version replaced gi:23097379.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GUCU
Center Clone name: CH230-365P7
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 125022 bases at least Q40
Consensus quality: 128107 bases at least Q30
Consensus quality: 129968 bases at least Q20
Estimated insert size: 130783; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 47440: contig of 47440 bp in length
* 47441 47540: gap of unknown length
* 47541 230268: contig of 182728 bp in length
* 230269 230368: gap of unknown length
* 230369 254735: contig of 24367 bp in length
* 254736 254835: gap of unknown length
* 254836 275940: contig of 21105 bp in length
* 275941 276040: gap of unknown length
* 276041 277352: contig of 1212 bp in length
* 277353 278493: contig of 1141 bp in length
* 278494 282220: contig of 3627 bp in length
* 282221 282320: gap of unknown length
* 282321 284049: contig of 1729 bp in length
* 284050 284149: gap of unknown length
* 284150 286208: contig of 2059 bp in length.
Location/Qualifiers
1..286208
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-365P7"
6333..7149
/note="clone_boundary
clone_end:77
site:
end_sequence:B2116846"
6368..7302
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:B2116847"
47541..48868
/note="wgs_end_extension
clone_end:Sp6"
230369..232081
/note="wgs_end_extension
clone_end:Sp6"
ORIGIN
Query Match 11.3%; Score 111.2; DB 2; Length 286208;
Best Local Similarity 50.3%; Pred.No. 6.1e-06;
Matches 300; Conservative 0; Mismatches 293; Indels 3; Gaps 1;
QY 53 ATTTTATTTTACCTCGTCCTAGATTCTGTAATCTATTCATTTCATTTCA 112
Db 229419 AATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 229478
QY 113 ATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 172

D	B		229479	TATTTTTTTTATTCTTTTTTTTAAATAAAAAAATTTCTAAATTTTATTTTTTTATTT	229538
Q	Y		173	AAGAAATTTTAATTAATTTTATATATTTAGATAAAATTCCTAAACACTTTTACTTTTTT	232
D	B		229539	AAAAATTTTTT---TTTTTTTTGAATGTTGTAATATTTATTTTATCATTTTTTTTTATTTT	229595
Q	Y		233	TTAAAAAGAATTTCAATTCGGTTTTTTCTTAATTTAGTTTTTAATCTATACATAATTATAA	292
D	B		229596	TTATATTTTTTTTTTAAATATTTTTTTTTTTTTTTTTTTTTTAAATATTATATATTTTTTT	229655
Q	Y		293	AAATCTGATCGATTAGTGGTGCMAAGTCAAGTCACATCGAATTTTTTGTTCGAGAAAA	352
D	B		229656	TTTTTTTTTATTTATTTTTTTTAAATTTATTTTTTTAAATTTATTTTTTTTATTAATAT	229715
Q	Y		353	AAATAAAATTAACAACATTTTTTCGATTAATTTATTTATATATATAATAATAAACACAT	412
D	B		229716	AATTATGTTTTTTTTTTTATTTTAAATATTTTTTTTTTTTTTTTTTTTTTAAATTTTTTTT	229775
Q	Y		413	TTTTTAAATGTTGCAATAATTTTTTTAATTAABATTTTCAGCACACAATTTACACTCT	472
D	B		229776	TTTTTGTGTTATTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTATTTTTTTTTTTATTTT	229835
Q	Y		473	CATCATAAATTTAATCTTATTACCATAATTAATAATTTGTGAGGACAAATTTTTTTAATC	532
D	B		229836	TTTATATTTTTTTTTTTATTT	229895
Q	Y		533	TCACCCCATTAAGCATATTTAATTTTTTGTTCGATACTCTCTTATTTTCACTCCCTAAC	592
D	B		229896	TATATTTTTTTAAATTTATTTATTTTATTTTAAATTAATTTTTTTTTTTTTTTTTTTA	229955
Q	Y		593	ATTAATCATTAACCCCAATTTTGAACCTGTTAATTTCTTAACCTTTATTCACATATGT	648
D	B		229956	TTTTTTTTTTTATTTTATTTATATTTTTTTTTTTTATTTCTTTAATTTCTTTTTTTTTT	230011

RESULT 8
 AEO14832/c
 LOCUS
 DEFINITION
 AEO14832
 ACCESSION
 AEO14832 AE014185
 VERSION
 AEO14832.1 GI:23495046
 SOURCE
 Plasmodium falciparum 3D7
 ORGANISM
 Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE
 1 (bases 1 to 258658)
 Gardner,M.J., Hall,N., Nelson,E., White,O., Berriman,M., Hyman,R.W.,
 Carlton,J.M., Pain,A., Fung,K.E., Bowman,S., Paulsen,I.T.,
 James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
 Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
 Anguoli,S., Pettea,M., Allen,J., Selengut,J., Haft,D.,
 Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,
 Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
 Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
 Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
 Barrell,B.
 TITLE
 Genome sequence of the human malaria parasite Plasmodium falciparum
 JOURNAL
 Nature 419 (6906), 498-511 (2002)
 PUBMED
 12368864
 REFERENCE
 2 (bases 1 to 258658)
 Gardner,M.J.
 AUTHORS
 Direct Submission
 TITLE
 Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
 JOURNAL
 Medical Center Dr., Rockville, MD 20850, USA
 FEATURES
 Location/Qualifiers
 1..258658
 /organism="Plasmodium falciparum 3D7"
 /mol_type="genomic DNA"
 /isoType="3D7"
 /db_xref="taxon:36329"
 /chromosome="10"
 -92..>2542
 source
 gene

mRNA	/locus_tag="PF10_0181" join<92..2026,2351..>2542) /locus_tag="PF10_0181"
CDS	join[92..2026,2351..2542) /locus_tag="PF10_0181" /codon_start=1 /product="hypothetical protein" /protein_id="AA055379.1" /db_xref="GI:23495047" /translation="MVHIQESSYNNQKQNDK ENNIYLRYSQDMKKINDCLNYNKKSSTLK NKKKDPYNNFKRHAYGADNYSFDVNNKKY SHYKRYKKKKEMDI EYKNKTWYDKYKEMK KDI SFLLSNDNIYRNHKIYIYNLEKYTD EYVSHNNYDI TEDEKNVIKINNPIYHST I SHVFFSKQITWKVFNHNLCAPLSPKEV KILFTFKNGSVI IWFENFNHOKNDNF INKI ROHNKWNDOCPNNKKNLKKCKEKKNNYS KSNELREMKKKKKKSLKSLSYEOKDKEN LSQSIOLDVHEMMDIATNLFKLSKEIAKH DELDIPEYFNWYKVEHAWFEIYKYMELPER FYTYRIMLLLFTHVLALLNDFLVV" complement(261..351) /notes="AT rich" /rpt_type=tandem 477..525 /rpt_type=tandem /rpt_unit=" (caaa)n" 581..691 /rpt_type=tandem /rpt_unit=" (a)n" complement(732..790) /note="AT rich" /rpt_type=tandem complement(838..914) /notes="AT rich" /rpt_type=tandem complement(1209..1246) /notes="AT rich" /rpt_type=tandem 1477..1623 /rpt_type=tandem /rpt_unit=" (taaaaa)n" 1484..1524 /rpt_type=tandem /rpt_unit=" (a)n" 1640..1663 /rpt_type=tandem /rpt_unit=" (a)n" 2034..2075 /rpt_type=tandem /rpt_unit=" (a)n" 2079..2233 /rpt_type=tandem 2092..2197 2093..2200 /rpt_type=tandem /rpt_unit=" (ttaaa)n" complement(2222..2273) /rpt_type=tandem 2222..2269 /rpt_type=tandem /rpt_type=tandem complement(2278..2298) /note="AT rich" /rpt_type=tandem complement(2308..2364) /note="AT rich" /rpt_type=tandem complement(2575..2644) /rpt_type=tandem /rpt_unit=" (ta)n"

```
repeat_region 2575..2628
gene /rpt_type=tandem
      complement (<2637..>4787)
mRNA /locus_tag="PF10_0182"
      complement (<2637..>4787)
CDS /locus_tag="PF10_0182"
      complement (2637..4787)
      /locus_tag="PF10_0182"
      /codon_start=1
      /product="hypothetical protein"
      /protein_id="AA35380.1"
      /db_xref="GI:23495048"
      /translation="MRIFGSLKRWYYSPYEYIILCNEQSKRRKNFSTMLYNTSEL
QNMMKNLENMGLISYGIKKSKYGVNKKMFHYNNLIKKFYCNFTFEDILILQ
SYALCKERNFEYSLLSKLLVIKDMNEQYQYDNIYKIILASQOLANSYDEII
TMFLKIIKNLHIGLKKISIIILSLTKLIYDQLLDTSSIIYVNNFEHMKCNVSH
IISAYSQIINIPLQVLIYENIKYMDLSIYNTLIOIKPILOTVDNYSLSY
KTQDDYCTLQNPKNKSYSSPEDNLFQKKKKKVDKGNINYESSNVCDDKYIKEVLQ
KNSYNNNTYDHSNVKRVHTYQSKETYPHTIPLVDYIEKHTKCTTNNINNDKQKK
NDENNVISFSSKSHLSDIYCHLNDTLKINNINKEQNNLKKKKLILNIPILFS
KYNNCLAFUSFQVLQACTYKELNYFNVIYIKRLLPYLANKIQNKIISPECVLL
ECFTVLPIYDNNIDETIILNTLETVITFYNYLCRIILQCFQHLTIYNDILSIDC
IIFKNKAPERYCNLDQLFLHFYEKNSDEQEMINYLKQIEQNNHILQSENSI
QODTHQDSTNNMTSMNLYNMDKKLIIVKYNKVKCFNEKEKNIYEHTNVHDDKQD
ENDGTPVHNSNDVRKVSDDIYFNTINEQKNI"
      complement (2821..2848)
      /note="AT rich"
repeat_region /rpt_type=tandem
      complement (3335..3365)
      /note="AT rich"
repeat_region /rpt_type=tandem
      complement (3453..3531)
      /note="AT rich"
      /rpt_type=tandem
      complement (3861..3880)
      /rpt_type=tandem
      complement (4042..4111)
      /note="AT rich"
      /rpt_type=tandem
      complement (4275..4301)
      /note="AT rich"
repeat_region /rpt_type=tandem
      complement (4348..4369)
      /note="AT rich"
      /rpt_type=tandem
      complement (4794..4992)
      /note="AT rich"
      /rpt_type=tandem
      complement (4833..4872)
      /rpt_type=tandem
      /rpt_unit="(a)n"
      4837..4872
      /rpt_type=tandem
      5124..5152
repeat_region /rpt_type=tandem
      /rpt_unit="(a)n"
      complement (5154..5178)
      /rpt_type=tandem
      /rpt_unit="(a)n"
      5275..5331
      /rpt_type=tandem
      5280..5323
repeat_region /rpt_type=tandem
      complement (5332..5357)
      /note="AT rich"
      /rpt_type=tandem
      complement (5378..5402)
      /note="AT rich"
      /rpt_type=tandem
      complement (5451..5484)
      /rpt_type=tandem
```

```
repeat_region /rpt_unit="(ta)n"
      complement (5495..5532)
      /note="AT rich"
repeat_region /rpt_type=tandem
      complement (5578..5604)
      /note="AT rich"
repeat_region /rpt_type=tandem
      complement (5670..5689)
      /rpt_type=tandem
      /rpt_unit="(a)n"
      complement (5749..5838)
      /rpt_type=tandem
      5753..5838
repeat_region /rpt_type=tandem
      complement (5876..5900)
      /note="AT rich"
      /rpt_type=tandem

Query Match 11.2%; Score 110.2; DB 3; Length 258658;
Best Local Similarity 51.4%; Pred. No. 8.3e-06;
Matches 279; Conservative 0; Mismatches 263; Indels 1; Gaps 1;

QY 83 TAAATACTATTCGATTCTCTCATTTTCATTATTTTAAATTTTATTTATTTATTTGGAT 142
DB 187113 TTAATTTTATATAATTAATAGTATATTTATTCATTTTAAATTAATATAATTAATTT 187054

QY 143 AAAAAATCTAATACCTTTACTTTTTTAAAAAGAAATTTTAAATTTATTTTATTTATTT 202
DB 187053 AATAATATAATTTATTTAAATTTATTTTGTGTTAAATTAATTAATTAATTAATA 186994

QY 203 AGATAAAATTCCTAATACCTTTACTTTTTTTTAAAGAAATTTCAATTCGGTTTTTCTT 262
DB 186993 TTTTATTTTAAATTTTATTTTAAATTTTATTTATTTATTAATTAATTTATTTAT 186934

QY 263 AATTTAGTTTTAACTTACTATACTAAATATAAAATTTCTGATCGATTAGTGTGTCATA 322
DB 186933 ATTTTATTTATTTAAATAAATATTTATTTGTTATTTATTTATTTATTTATTTAT 186874

QY 323 GTCAAGTCACATGAATTTTGTGGAGAAAAATAAAAAATAAACACATTTTTCGATTAAAT 382
DB 186873 TATTATTTATTTAAATAAATAATATTTGTTATTTATTTATTTATTTATTTATTT 186814

QY 383 TTATTATATATAATAATAATAACACATTTTATTTTAAATGTTGCAATAATAATTTTAA 442
DB 186813 TTATTATATTTAAATAAATAATTTATTTGTTATTTATTTATTTATTTATTTATTT 186754

QY 443 ATTAAATTTTCAGCACACAATTTACACTCTATC-ATTAAATTTAATCTTATTAACATAA 501
DB 186753 ATTATTTTAAATAATAATAATATGTTATTTATTTATTTATTTATTTATTTATTTAT 186694

QY 502 TTTAAATTTGTGAGGACAATTTATTTTAAATCTCACCTCCATTAATGTCATATTTAAAT 561
DB 186693 TTTAAATATAATATGTTATTTATTTATTTTATTTATTTAAATTAATTAATGCTATA 186634

QY 562 TTTGTTGATACCTTTCTTATTTTCCTCACTCACTCACTCACTCACTCACTCACTCACT 621
DB 186633 TATTGTTATTTATCTTTTATTAATAATCAATGTTTAAATTAATAATAATAATAATA 186574

QY 622 ATA 624
DB 186573 TTA 186571

RESULT 9
CR847965/c 191191 bp DNA linear HTG 11-OCT-2004
LOCUS Danio rerio clone CH211-223F21, *** SEQUENCING IN PROGRESS ***, 9
DEFINITION unorderd pieces.
ACCESSION CR847965
VERSION CR847965.2 GI:54034964
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Danio rerio (zebrafish)
```


KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

HTG.
 Plasmodium falciparum 3D7
 Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 348174)
 Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,
 Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,
 Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C.,
 Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,
 Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,
 Felwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,
 Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,
 Horrocks, P., Humphray, S., Jageis, K., James, K.D., Johnson, D.,
 Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N.,
 Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,
 Moulle, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A.,
 Roubiniwitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M.,
 Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,
 Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, I.,
 Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and
 Barrell, B.G.
 Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
 Nature 419 (6906), 527-531 (2002)
 22255708
 12368867

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 348174)
 Cherevach, I., Davis, P., Goodhead, I., Stevens, K., Mungall, K.,
 Berriman, M., Pain, A., Hall, N., Atkin, R., Chillingworth, C.,
 Doggett, J., Ormond, D., Sanders, M., Hayes, R., Hall, S., Quail, M. and
 Barrell, B.G.
 Unpublished

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

3 (bases 1 to 348174)
 Cherevach, I., Davis, P., Goodhead, I., Stevens, K., Mungall, K.,
 Berriman, M., Pain, A., Hall, N., Atkin, R., Chillingworth, C.,
 Doggett, J., Ormond, D., Sanders, M., Hayes, R., Hall, S., Quail, M. and
 Barrell, B.G.
 Direct Submission

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (20-SEP-2002) P. falciparum Genome Sequencing Consortium,
 The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,
 Hinxton, Cambridge CB10 1SA, UK
 4 (bases 1 to 348174)
 Cherevach, I., Davis, P., Goodhead, I., Stevens, K., Mungall, K.,
 Berry, A.E., Berriman, M., RA Pain, A., Hall, N., Atkin, R.,
 Chillingworth, C., Doggett, J., Ormond, D., Sanders, M., Hayes, R.,
 Hall, S., Quail, M. and Barrell, B.G.
 Direct Submission

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (26-MAR-2004) P. falciparum Genome Sequencing Consortium,
 The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,
 Hinxton, Cambridge CB10 1SA, UK
 For more information about this sequence or the Malaria Project,
 see http://www.sanger.ac.uk/Projects/P_falciparum.
 Location/Qualifiers

FEATURES
source

1..348174
 /organism="plasmodium falciparum 3D7"
 /mol_type="genomic DNA"
 /isolate="3D7"
 /db_xref="taxon:36329"
 /chromosome="6"
 complement(847..1143)
 /locus_tag="PFF0415C"
 complement(847..1143)
 /locus_tag="PFF0415C"
 /note="MAL6P1.87"
 /codon_start=1
 /product="hypothetical protein, conserved"
 /protein_id="CAG25326.1"
 /db_xref="GI:46361039"
 /translation="NNFSKILFLKKHQTIIIRNVFSLSSKNGNMFKRAASTENKQNG
 EOKDOENVYKANC PKGIRIPIMRYFYGLIFLMAFPVPTQSLYETNKYKVS"
 complement(join(2408..3006,3295..3403))
 /locus_tag="PFF0420C"
 complement(join(2408..3006,3295..3403))
 /locus_tag="PFF0420C"

gene
CDS

/EC_number="3.4.25.1"
 /note="MAL6P1.88"
 /codon_start=1
 /product="proteasome subunit alpha type 2, putative"
 /protein_id="CAG25327.1"
 /db_xref="GI:46361040"
 /translation="MADGESFSLTTFPSPTGKLVQIEVALNRVSSSSPALGRKNGV
 IATEKSPNELIEENSIFKIQISEHIGIVYAGMPGPRVLLKRAEKAIRYSQYQ
 SEILKEVLKIIASIVQSFOTGTGVRPGLSLICGVQVYHLLQIDSPSGYFNWMA
 TCVGKDYNNMSFLEKRYNKQIDIEDAHTAILTKSEYEGVLEKNEKIBIGVAYDNKP
 FKILQNEIKDYLIBIE"
 join(5857..5887,6084..6904)
 /locus_tag="PFF0425W"
 join(5857..5887,6084..6904)
 /locus_tag="PFF0425W"
 /note="MAL6P1.89"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAG25328.1"
 /db_xref="GI:46361041"
 /translation="MEDSSLENKSTEGNLESEYNEENQNEENESVENESVE
 NESENVNEENVEENKNEKNKNMNYQENGNNNNNNNKNDKLNKSPHNNNEL
 KNNLNNLNNNNNNNNLGTKNLENEGFSYKNOEDLVNDNKYSNEKYGDM
 THSKNSPKNNEHNMIDEQKSSSHEKKNRYDNLVFPQSDKRNLFQGHNKKTFFP
 SLSPDVLIPGKEHENITMHPNPEKFGIKYNNGLKPSGEGYLLIIPHISHNSIAPKIK
 KKKRCC"
 join(9019..10528,10710..10831)
 /locus_tag="PFF0430W"
 join(9019..10528,10710..10831)
 /locus_tag="PFF0430W"
 /note="MAL6P1.90"
 /codon_start=1
 /product="chaperone, putative"
 /protein_id="CAG25329.1"
 /db_xref="GI:46361042"
 /translation="MSIHLLKAKADSLRSTNVMTNMNASKMGYIILKSNLKPCKSYK
 MLVSSGAIKITKQGNVLNEMIQHTASMLSRICSSIDETLGGSSNLIVATSLI
 YLSEKYLINHPRIITQGFDAIKGILFELDSMKIPVNIENEDKELLYNVKTCI
 RTKIPICLADLADLVESIKIYKPTQIDLHMEIMDKNNMINTKLVRGWLDDH
 GRHPNMNKLTKCFLVNLSEYSEKVSFSSFYNAEDRDLKVESRKTDDKVK
 KTELKTLVEKPKETWEIYNAFVNKGIDIPISLDLLAKENIHALRRIKRNLERI
 VLCCGNCNNVYDLTEEDGVAGLVYISINDEKTYETEEVQNPSCSTIFQAPNDY
 TIKQIDAIRGLRSKINVIDDKVLSGAGSFEIMAYCKDKDEKKIKGKPKALDIY
 ANSLNIPKVLLENSGLDIHQTLFNVIDKYNEDRSEPLGLDLDTGPTIAHLKGIYDN
 YCVKKEILSIATISQIILLVDEITRAGKSMGEK"
 13263..14507
 /locus_tag="PFF0435W"
 13263..14507
 /locus_tag="PFF0435W"
 /EC_number="2.6.1.13"
 /note="MAL6P1.91"
 /codon_start=1
 /product="ornithine aminotransferase"
 /protein_id="CAG25330.1"
 /db_xref="GI:46361043"
 /translation="MDFVKELKSSQDYNNELTYGAHYNDYPIPVVLKRGKGVFVVDIE
 DRYVDYFLSAYSSVNOGCHPDILNAMIQAOKLTICSRAPFSDSLGVCERYLTNLFG
 YKVLMMNTGAEASETA YKLCRWGEYVKI PENSAKIIVCNNFSGRITLGVASSTD
 KKKNPGFPVFNELKVPYDDLAELEKELODPNVCAFIPEVQAGVIVPSDSYFPFG
 VASLCKNLVFLVADEVOTGLRGTGKLLCTHHYGVKVPDILLGKALSGHYFISAILA
 NDDVNLKLPGEHSGSTYGNPLAAACIVALKVLINEKLECNADKLGAFFLQNLKEQL
 KDSKVVREVGRGKLLCAIEFKNDLVNWDICLKFKENGLITRSVHDKTVRLTPPLCIT
 KEQLDECTEIIIVTKVFFDDNL"
 join(15915..18857,19043..19087)
 /locus_tag="PFF0440W"
 join(15915..18857,19043..19087)
 /locus_tag="PFF0440W"
 /note="MAL6P1.92"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAG25331.1"
 /db_xref="GI:46361044"
 /translation="MNYGICDTLKKKKSSSYLLETNETVKKLMMNEGEDMNNFMINKK

* 125994 169056: contig of 43063 bp in length.

```

FEATURES
    source
        1. .169056
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone="CH211-235C15"
            /clone_lib="CHORI-211"
            1. .10976
                /note="assembly fragment:00034"
                fragment_chain:1
                clone_end:T7
                vector_side:left"
            11077..95174
                /note="assembly fragment:01332"
                fragment_chain:1"
            95275..106934
                /note="assembly fragment:00170"
                fragment_chain:1"
            107035..125893
                /note="assembly fragment:00340"
                fragment_chain:1"
            125994..169056
                /note="assembly fragment:00675"
                fragment_chain:1
                clone_end:SP6
                vector_side:right"

ORIGIN
    Query Match      11.1%; Score 109.6; DB 5; Length 169056;
    Best Local Similarity 53.8%; Pred. No. 1e-05;
    Matches 300; Conservative 0; Mismatches 244; Indels 14; Gaps 3;

QY  53 ATTTTATTTTACCTCGCTAGATTCGTAATACTATGTCATTATCTCATTTCA 112
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  62305 ATTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 62246
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  113 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 172
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  62245 ATTGTATATAAATCTATTTTGTGTATGTTAATTTCA 62188
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  173 AGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 232
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  62187 TTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 62128
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  233 TTAATAAGAAATTTCA-----ATTGCGTTTTTCTTAATTAGTTTAAATCTATA 282
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  62127 TTTAATGATTTTATTTATTTAAGTTTAAATTAATGAATTTATTTATTTATTTA 62068
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  283 CTAATTTAAAAAATCTGATCGGATPAGTGGTGCAAGTCAAGTCACATGAATTTTG 342
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  62067 TTCTTTATAAAGTTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 62008
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  343 TTGGAGAAAAATAAAAAATAAACACATTTTTCGATTATTTATTTATTTATTTA 402
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  62007 TTATTTTATAATTTAATGTTTATTTATTTATTTATTTATTTATTTATTTAT 61948
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  403 TAAACACATTTTATTTAATGTTGCAATAATTTTAAATTTAAATTTAAATTTAA 460
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  61947 TTATTTCTTCATATAATAACAATTTTATTTTGTGTTATTTTATTTTATTTA 61888
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  461 CAATTACACTCTCATCAATTTAAATCTTATTTACCAATAAATAAATTTGTGAG 520
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  61887 TTATTTTATAAATTTGTTTGAATTTATTTCTTCATAATAACTTTATTTTAT 61828
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  521 TATTTTATTTATCTACCCCTCAATTAATGATTTATTTATTTTGTTCGATCTTAT 580
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  61827 GTATTTTATTTGTTATGTTTAAATTTAATTTAATTTATTTTATTTTATTTAT 61768
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  581 TTCACCTCTCAACATAAAT 598
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  61767 TTCTGACTTAATTAATTA 61750
DB  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 12

BX247875/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 213093)

Ellwood, M.

Direct Submission

Submitted (25-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Oct 21, 2003 this sequence version replaced gi:37518258.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived

zebrafish PUC subclones occasionally display inconsistency over the

length of mononucleotide A/T runs and conserved TA repeats. Where

this is found the longest good quality representation will be

submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat

discovery system (Zhirong Bao and Sean Eddy, submitted), and those

beginning 'drr' were identified by Rick Waterman (Stephen Johnson

lab, WashU). For further information see

http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml

CH211-201N18 is from a CHORI-211 BAC library

VECTOR: pTARBAC2.1.

FEATURES

Location/Qualifiers

1..213093

/organism="Danio rerio"

/mol_type="genomic DNA"

/db_xref="taxon:7955"

/clone="CH211-201N18"

/clone_lib="CHORI-211"

ORIGIN

Query Match 11.1%; Score 109.6; DB 5; Length 213093;

Best Local Similarity 53.8%; Pred. No. 1e-05;

Matches 300; Conservative 0; Mismatches 244; Indels 14; Gaps 3;

QY 53 ATTTTATTTTACCTCGCTAGATTCGTAATACTATGTCATTATCTCATTTCA 112

DB 163805 ATTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 163746

Lorensuhewa, L., Loulsegged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Manung, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muniadasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackemeleh, O., Okwuonu, G., Olarnpunaagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfaunkoch, C., Piopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajis, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soëa, J., Stealmle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished
2 (bases 1 to 246611)

Worley, K.C.

Direct Submission

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 246611)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:24819049.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GMFA

Center clone name: CH230-61K4

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 217670 bases at least Q40

Consensus quality: 219537 bases at least Q30

Consensus quality: 220430 bases at least Q20

Estimated insert size: 231456; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

- * NOTE: Estimated insert size may differ from sequence length
- * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 4 contigs. The true order of the pieces
- * is not known and their order in this sequence record is

- * arbitrary. Gaps between the contigs are represented as
- * runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will
- * be preserved

- * 1 241059: contig of 241059 bp in length
- * 241060 241159: gap of unknown length
- * 241160 242345: contig of 1186 bp in length
- * 242346 242445: gap of unknown length
- * 242446 243601: contig of 1156 bp in length
- * 243602 243701: gap of unknown length
- * 243702 246611: contig of 2910 bp in length.

FEATURES

source

1. 246611

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clones="CH230-61K4"

1. 1746

/note="wgs_contig"

misc_feature

ORIGIN

Query Match 11.1%; Score 109.2; DB 2; Length 246611;

Best Local Similarity 51.3%; Pred. No. 1.1e-05;

Matches 307; Conservative 0; Mismatches 284; Indels 8; Gaps 2;

Qy	52	AAATTTTATTTTACCTCCTCGTAGATTCGTAATACTATTGCAATTTATCATTTTCAAT	111
Db	215405	ATTTTTCGTTATTTTATTTTATATATTTTATTTTATTTTATTTTATTTTATTTTAT	215346
Qy	112	TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	171
Db	215345	TTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	215286
Qy	172	AAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT	226
Db	215285	ATAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	215226
Qy	227	TTTTTTTTTAAAGAAATTTCAATTCGTTTTTCTTAATTTAGTTTAAATTTCTATAC	286
Db	215225	TAAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	215166
Qy	287	TTATATAAATTTCTGATCGGATTTAGTGTGTCGAAGTCACATGCAATTTTGTGG	346
Db	215165	TTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	215106
Qy	347	AGAAAAATATAAATAAACAATTTTTCGATTTAATTTATTTATATATATAATAATAA	406
Db	215105	ATTTTATTTATTAATAGTATATTTT---ATATATTTTATTTTATTTTATTTTATTT	215049
Qy	407	CACATTTTATTTAAATTTGTCGAATAATTTTAAATTTAAATTTTCAACACAAATTA	466
Db	215048	TTATTTTAAATTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	214989
Qy	467	CACCTCATCATTAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTT	526
Db	214988	ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	214929
Qy	527	TTAACTCTCACCTCCATTAATGATATTTAATTTTTCGTCGATACCTTCTTTTCACT	586
Db	214928	TTATATTAATTAATTTTAAATTAATTAATTAATTTTATTTTATTTTATTTTATTT	214869
Qy	587	CCTAACATTTAATCATTAACCCCAATTTTGAATTTTGAATTTTCTTAACATTTTCA	645
Db	214868	TTTAAATTTTATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	214810

Search completed: March 15, 2005, 01:02:23

Job time : 4305 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 23:13:46 ; Search time 209 Seconds
(without alignments)

7711.639 Million cell updates/sec

Title: US-10-089-543-2 COPY 449 1433

Perfect score: 985
Sequence: 1 gatgtgaccagtggtggag.....tgagagaaatccttcacatc 985

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 24055568

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

.....C.....F.....

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

```
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C 1	101.8	10.3	612	4	US-09-902-540-1357	Sequence 1357, App
C 2	90.2	9.2	59519	4	US-09-949-016-13504	Sequence 13504, A
C 3	88.8	9.0	55886	4	US-09-949-016-15129	Sequence 15129, A
C 4	88.2	9.0	205044	4	US-09-949-016-15851	Sequence 15851, A
C 5	88.2	9.0	205044	4	US-09-949-016-15852	Sequence 15852, A
C 6	88.2	9.0	205044	4	US-09-949-016-15853	Sequence 15853, A
C 7	88.2	9.0	223471	4	US-09-949-016-12387	Sequence 12387, A
C 8	88.2	9.0	223471	4	US-09-949-016-12724	Sequence 12724, A
C 9	88.2	9.0	223471	4	US-09-949-016-12725	Sequence 12725, A
C 10	87.8	8.9	61178	4	US-09-949-016-17369	Sequence 17369, A
C 11	87	8.8	19124	2	US-08-487-826B-13	Sequence 13, App1
C 12	85.8	8.7	187169	4	US-09-949-016-12776	Sequence 12776, A
C 13	85.8	8.7	191589	4	US-09-949-016-15940	Sequence 15940, A
C 14	84.4	8.6	1039	4	US-09-902-540-1280	Sequence 1280, App
C 15	84.2	8.5	263693	4	US-09-949-016-12386	Sequence 12386, A
C 16	84.2	8.5	263694	4	US-09-949-016-16915	Sequence 16915, A
C 17	84	8.5	20674	3	US-09-641-638-651	Sequence 651, App
C 18	84	8.5	20674	4	US-10-170-097-651	Sequence 651, App
C 19	83.8	8.5	147382	4	US-09-949-016-14624	Sequence 14624, A
C 20	83.6	8.5	601	4	US-09-949-016-30530	Sequence 30530, A
C 21	83.6	8.5	601	4	US-09-949-016-30531	Sequence 30531, A
C 22	83.6	8.5	601	4	US-09-949-016-37149	Sequence 37149, A
C 23	83.6	8.5	601	4	US-09-949-016-37150	Sequence 37150, A
C 24	83.6	8.5	601	4	US-09-949-016-37163	Sequence 37163, A
C 25	83.6	8.5	601	4	US-09-949-016-37164	Sequence 37164, A
C 26	83.6	8.5	601	4	US-09-949-016-145867	Sequence 145867, A
C 27	83.6	8.5	601	4	US-09-949-016-145868	Sequence 145868, A

ALIGNMENTS

RESULT 1

```

US-09-302-540-1357/C
/ Sequence 1357, Application US/0902540
/ Patent No. 6833447
/ GENERAL INFORMATION:
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902
/ CURRENT FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883
/ PRIOR FILING DATE: 2000-07-10
/ NUMBER OF SEQ ID NOS: 16825
/ SEQ ID NO 1357
/ LENGTH: 612
/ TYPE: DNA
/ ORGANISM: Myxococcus xanthus
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(612)
/ OTHER INFORMATION: unsure at all n l
US-09-302-540-1357

```

```
Query Match      10.3%; Score 101.8; DB 4; Length 612;
Best Local Similarity 48.1%; Pred. No. 2.6e-10;
Matches 286; Conservative 0; Mismatches 309; Indels 0; Gaps 0;
```

[illegible]

Db 171180 ATATATATAAATATATTTTATATATATTTTAAATATATATAAATATATATTTTATATA 171121
QY 519 ATTATTTTTTAATCTCACCTCCATTAAATGCAATATTTAAATTTT 563
Db 171120 TTTATATTTAAATATATAAATATATTTTATATATTTTATATTTT 171076

RESULT 8
US-09-949-016-12724/c
; Sequence 12724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12724
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724

Query Match 9.0%; Score 88.2; DB 4; Length 223471;
Best Local Similarity 50.5%; Pred. No. 1.8e-07;
Matches 265; Conservative 0; Mismatches 258; Indels 2; Gaps 2;
QY 40 TCAAGTGGCACCAATTTTTTATTTTACCTCCCTAGATTCGTAATACTATTGCAATTT 99
Db 171599 TCATGTTTCATAGATAGGGGTCAATTCCTCAITTTAAATATATTTCAAAATATTAATATATTT 171540
QY 100 ATCTCATTTTCATTATTTTAAATTTTATATTTTAAATTTTGGATATAAAATCTTAATCTTT 159
Db 171539 ATATTTAAATATATAAATATATTTTATATATTTTAAATATATTTTAAATATATATTTTAT 171480
QY 160 ACTTTTTTTTAAAGAAATTTATTTTAAATTTTATATTTTAAATTTTAGATAAAATCTTAATA 219
Db 171479 ATATTTTAAATATATAAATATATTTTATATATTTTAAATATATTTTAAATATATATTTTAT 171420
QY 220 CTTTACTTTTTTTTAAAGAA-ATTTCAATTCGGTTTTTCTTAATTTTAGTTTAAATTC 278
Db 171419 ATATATTTTAAATATATAAATATATATTTTATATATTTTAAATATATATATTTTAT 171360
QY 279 TATACATAATATAAATTTCTGATCGGATAGTGCTCAAGTCAAGTCACATGAAT 338
Db 171359 TTATATATTTTAAATATATAAATATATTTTATATATTTTAAATATATATATTTTAT 171300
QY 339 TTTGTTGGAGAAAAATAAATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 398
Db 171299 TTTTATATTTTAAATATATAAATATATTTTATATATTTTAAATATATA 171241
QY 399 AATATAAACACATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 458
Db 171240 TATTTTATATTTTAAATATATAAATATATAAATATATAAATATATAAATATATAAATTTTAA 171181
QY 459 AACAAATACACTCTCATCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 518
Db 171180 ATATATAAATATATTTTATATATTTTAAATATATAAATATATAAATATATAAATTTTAA 171121
QY 519 ATTATTTTTTAATCTCACCTCCATTAAATGCAATTTTAAATTTT 563

Db 171120 TTTATATTTAAATATATAAATATATTTTATATATTTTATATATTTTATATTTT 171076

RESULT 9
US-09-949-016-12725/c
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12725
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725

Query Match 9.0%; Score 88.2; DB 4; Length 223471;
Best Local Similarity 50.5%; Pred. No. 1.8e-07;
Matches 265; Conservative 0; Mismatches 258; Indels 2; Gaps 2;
QY 40 TCAAGTGGCACCAATTTTTTATTTTACCTCCCTAGATTCGTAATACTATTGCAATTT 99
Db 171599 TCATGTTTCATAGATAGGGGTCAATTCCTCAITTTAAATATATTTCAAAATATTAATATATTT 171540
QY 100 ATCTCATTTTCATTATTTTAAATTTTATATTTTAAATTTTGGATATAAAATCTTAATCTTT 159
Db 171539 ATATTTAAATATATAAATATATTTTATATATTTTAAATATATATATTTTAAATATATATTTTAT 171480
QY 160 ACTTTTTTTTAAAGAAATTTATTTTAAATTTTATATTTTAAATTTTAGATAAAATCTTAATA 219
Db 171479 ATATTTTAAATATATAAATATATTTTATATATTTTAAATATATTTTAAATATATATTTTAT 171420
QY 220 CTTTACTTTTTTTTAAAGAA-ATTTCAATTCGGTTTTTCTTAATTTTAGTTTAAATTC 278
Db 171419 ATATATTTTAAATATATAAATATATATTTTATATATTTTAAATATATAAATATATAAATTTT 171360
QY 279 TATACATAATATAAATTTCTGATCGGATAGTGCTCAAGTCAAGTCACATGAAT 338
Db 171359 TTATATATTTTAAATATATAAATATATTTTATATATTTTAAATATATATATTTTAT 171300
QY 339 TTTGTTGGAGAAAAATAAATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 398
Db 171299 TTTTATATTTTAAATATATAAATATATAAATATATAAATATATAAATTTTAAATTTTAA 171241
QY 399 AATATAAACACATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 458
Db 171240 TATTTTATATTTTAAATATATAAATATATAAATATATAAATATATAAATTTTAAATTTTAA 171181
QY 459 AACAAATACACTCTCATCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 518
Db 171180 ATATATAAATATATTTTATATATTTTAAATATATAAATATATAAATATATAAATTTTAA 171121
QY 519 ATTATTTTTTAATCTCACCTCCATTAAATGCAATTTTAAATTTT 563
Db 171120 TTTATATTTTAAATATATAAATATATTTTATATATTTTATATATTTTATATTTT 171076

Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSES: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match	8.8%;	Score 87;	DB 2;	Length 19124;
Best Local Similarity	50.1%;	Pred. No. 2.2e-07;		
Matches 252;	Conservative	0;	Mismatches 240;	Indels 11; Gaps 1;
QY	58	TTATTTTACCTCGCTAGATTGTAATACATATTGCAATTTATCTCATTTTCATTTATTTA	117	
Db	15878	TTTTTTTTATTATTATTTTTTTTTTTTATTAAATTTTTTTTATTTATTTATTTATTTTTTTTTT	15815	
QY	118	TTTAAATTTATTTATATTATTTGGATAAAAATTCCTAATCTTTACTTTTTTTTTTAAAGAA	177	
Db	15818	TTTTAAATAAATTTTTTTTATTTATGTATATATTTTTTTTTTTTAAACATTTTTTTTAAATTTT	15755	
QY	178	TTTTATTAAATTTATTTATATATTTAGATAAAAATCTCAATCTTTTACTTTTTTTTTTAAA	237	
Db	15758	TTTTATTTTATGATATATATTTTTTATTTTAAATATATTTTTTTCTTTTTTTTTTGTTTTT	15691	
QY	238	AAGAAATTCAAATCGGTTTTTTCTTAAATTTAGTTTTTAAATCTATCTATTAATAAAAAAT	297	
Db	15698	TGATATATATTTTTTTTTTTTTTAAAGTTTTTTTTTTTCTCTTTCTGTTTTATTTTTTTT	15631	
QY	298	CTGATCGGAATTAGTGTGGTGTCAAAGTCAAGTCAATGAATTTGTGTGGAGAAAAATAA	357	
Db	15638	ATAATCATTTTTTTTTTATATAAAAAATTTTTTTTTTAAATTTTTTTTGTATAATCTTTTCAT	15571	
QY	358	AAATTTAAACACATTTTTTCGATTATTTATTTATATATATAATAATAACACATTTTTTAT	417	
Db	15578	TTTTTTATCTATCAAAATTTATATTTTTTATATATAATTTTTTATATTTTTTTTAAAAATTTT	15511	

RESULT 10

US-09-949-016-17369

Sequence 17369, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 17369

LENGTH: 61178

TYPE: DNA

ORGANISM: Human

US-09-949-016-17369

Query Match 8.9%; Score 87.4; DB 4; Length 61178;

Best Local Similarity 49.9%; Pred. No. 2.1e-07;

Matches 275; Conservative 0; Mismatches 271; Indels 5; Gaps 2;

Qy 98 TTATCTCATTTTCATTTATTTATTAATTTTATATTTTATATTTTGGATAAAATTCCTAACT 157

Db 21250 TAATCTTATCTTATATAAATCTTATATTTATATGCTCTTATATAAATATATCTTATATAT 21309

Qy 158 TTACTTTTTTTTAAAAAGAAATTTATTTAAATTTTATATTTTATATTTTAGATAAAAAATCTTAA 217

Db 21310 ATAACTTATCTTATATATAAATCTTATTTATATAAATAAATCTTATTTATATAATATAA 21369

Qy 218 TACTTTACTTTTTTTTTTAAAAAGAAATTTCAATGGCTTTTCTTAAATTTAGTTTAAAT 277

Db 21370 TCTTATCTTATATAAATAAATCTTATTTATATAAATAAATCTTATCTTATTTATATAAT 21429

Qy 278 CTATACTAAATTTATAAAATTTCTGATCGGATTAGTGTGGTGTCAAAAGTCAAGTCAATGAA 337

Db 21430 AATATCTTATATAAATAAATAAATCTTATTTATATAAATCTTATTTATATAAAT 21488

Qy 338 TTTTGTGGAGAAAAATAAAAAATTAACACATTTTTCATTAATTTATATATATATAA 397

Db 21489 ATAATCTTATCTTATATAAATAAATCTTATTTATATAAATAAATAAATCTTATCTTA 21548

Qy 398 TAATATAAACAATTTTTTATTAATGTTGCAATATAATTTTTTAAATTTAAAAATTTTCAGCA 457

Db 21549 TTATATAATAAATCTTATCTTATATAAATAAATAAATCTTATATAAATAAATCTTAT 21608

Qy 458 CAACAATTTACACTCTCATCAATTAATTAATCTTAT---TACATAAATTAATTTGGA 513

Db 21609 TATATAATAAATCTTATTTATATAAATCTTATTTATATAAATAAATAAATCTTATATTA 21668

Qy 514 GGACAATTTTTTAAATCTCACCCCTCAATTAATGCATATTAATTTTCTGTCGATAC 573

Db 21669 ATATAAATCTTATTTATATAAATAAATCTTATTTATATAAATAAATAAATCTTATATA 21728

Qy 574 TTCTTATTTCACTCCTCAACATTAATCAATTAACCAATTTTGAATCTGTATATAATTTCTTAA 633

Db 21729 TCTTATCTTATATAAATGTTATTTATATAAATCTTATCTTATATAAATCTTAT 21788

Qy 634 CTTATTTCACTA 644

Db 21789 ATTATTATATA 21799

RESULT 11
US-08-487-826B-13/c

Qy	418	TTAATGTTGCAATAATATATTTTTTAATTTAAATTTTCAGCACAAACAATTTACACTCTCATCA	477
Db	15518	CTCCITTTTTTTTTTTTTTTTTTTATTTTAAAT-----AAATTTTTTTTTTATAT	15470
Qy	478	TTAAATTTAATCTTATTAACCAATAATTTAAATTTGTGAGGACAAATTTTTTTTAAATCTCACCC	537
Db	15469	TTCAATTTTCTTTTTTTCATTTTAAATAAATTTGTTTTTATATTTCTTTTTTAAATATAAAT	15410
Qy	538	CTCCATTAATGCATATATTATTAAT	560
Db	15409	ACATATATATAAATAATATATATAT	15387

```

RESULT 12
US-09-949-016-12776/c
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(187169)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

```

				8.7%; Score 95.8; DB 4; Length 187169;	
				Best Local Similarity 51.0%; Pred. No. 4.9e-07;	
				Matches 281; Conservative 0; Mismatches 262; Indels 8; Gaps 3;	
Qy	90	TATTGCATTTATCTCATTCGATTAATTTTAAATATTTTATATATAT-TTGCGATAAAAAAT	148		
Db	1036	TATTATATATTTTATATTTTATATATATATATATATATTTTATATATATATATATATAT	977		
Qy	149	TCTAATACTTTCACTTTTTTTTAAAAAGAAATTTATTTAAATATTTTATATATTTTAGATAA	208		
Db	976	TATATATGCCATATATTTTTTATATATATATATTTTATATATATATATATATATATATTA	917		
Qy	209	AAATTCATAACTTTTACTTTTTTTTTAAAAAGAAATTTCAAATGGCTTTTTTCCTTAATTA	268		
Db	916	TATATTTTATATATATATATATATTTTATATATATATATATATATATATATATATATATA	857		
Qy	269	GTTTTAAATCTATACTAATTAATAAAAATTCGTATCGGAATAGTGTGGTGC AAAAGTC AAG	328		
Db	856	TTTTTATATATATATATATTTTATATATAT - - - - - ATA TTTTATATATTTTATATATTT	803		
Qy	329	TCATATGAATTTCTGTCGAGAAAAATAAAAAATTAACACACATTTTTCGATTAATTTATTA	388		
Db	802	TTATATATTTATATATAATATATATTTTATATATTTTATATATATATATATATATTTTTT	743		
Qy	389	TATATATAAATAATAAACACATTTTTTATTTAAATGGTCTCAATAATATTTTTTAAATAAA	448		
Db	742	TATATATATATATATTTTATTTTATATATATATATATATATATATATATATTTTTATGTA	683		
Qy	449	ATTTCCAGCACAACAAATTACACTCTCATTCATTPAAATTTTAATCTTATTACCATAATTTAAAT	508		

[illegible]

```

RESULT 13
US-09-949-016-15940/c
; Sequence 15940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15940
; LENGTH: 191569
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(191569)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15940

```

Query Match	8.7%;	Score 85.8;	DB 4;	Length 191569;
Best Local Similarity	51.0%;	Pred. No. 4.9e-07;		
Matches 281;	Conservative 0;	Mismatches 262;	Indels 8;	Gaps 3;
Qy	90	TATTGCATTTATCTCATTTCCATTTATTTTAAATATATTTATATATAT- TTGGATATAAAAT	148	
Db	1036	TATTATATATTTTATATTTTATATTTTATATATATATATATATTTTATATATCATATATAT	977	
Qy	149	TCATAACTTTACTTTTTTTTAAAGAAATTTTAAATTTTATTTATTTAGATAA	208	
Db	976	TATATATCCATATATTTTATATATTTTATATATTTTATATATTTATATATATATTA	917	
Qy	209	AAATTCATACTTACTTTTTTTTAAAGAAATTTCAATCGGTTTTTCTTAATTTA	268	
Db	916	TATATTTTATATATATTTATATATTTTATATATTTTATATATATTTTATATATA	857	
Qy	269	GTTTTAACTCTACTAATTTATAAAATTCCTGATCGGATTTAGTGTGGTCAAAAGTCAAG	328	
Db	856	TTTTTATATATATATATATTTTATATATTTAT- ----ATATTTTATATATTTTATATATT	803	
Qy	329	TCATCGAAATTTGGTGGAGAAAAATAAAAAATTTAAACACATTTTTCGATTAATTTATTA	388	
Db	802	TTATATATTTATATAATATATATTTTATATTTTATATATATATATATATTTTATATTTT	743	
Qy	389	TATATATAATAATATAAAACACATTTTATTTAAATGTGTCAATAATATTTTTTAAATTTAA	448	
Db	742	TATATATATTTATATATTTTATATTTTATATATATATATATATTTTATATATATTTTATGTA	683	
Qy	449	ATTTTCAGCACACAAATTTACACTCTCATTTAAATTTAATCTTATTACCAATAATTTAAAT	508	

682	TATTATATATTATATATATATTTTATATATATATATATATATTTTATATATATATATATATAT	623
Db		
509	TGTGAGGACAATATTTTTTAAATCCACCTCCCAATTAATGCAATATAT-TAAATTTTGGT	567
Qy		
622	TATATATATTTTATATATATATATAGTATATATATTTTATATATACTATATAATATATATAT	563
Db		
568	CGATCTCTTTATTTTTCACCTCCCTAACATTTAACTCAATTAACCCCAATTTTGAAGCTGTTATAAT	627
Qy		
562	TTATATATATATATATATTTTATATATATATTTTATATATATATATATATATATATATATTTTT	503
Db		
628	TCTTAACTTAT 638	
Qy		
502	ATATATTTTAT 492	
Db		

RESULT 14
US-09-902-540-1280/c
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkie, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

[illegible]

QY	472	TCATCATTTAAATTAATCTTATTACCATATAATTAATAATGTGAGGACAAATATTATTTTAAAT	531
Db	188	TTTTTTTTTATTTTTTTTTTTTAAATATTTTTTTTTTTTTTTTTATATATTTGTTTTTATTTT	129
QY	532	CTCACCTTCATTAATGACATATTATTAAATTTTTTGGTTCGATCTCTTATTTCACCTCTAA	591
Db	128	TTTTTTTATTAATTTGTTTTTTTTTATTTTTTTTTTTTTTTTTTATTTTGTGTTTTTAA	69
QY	592	CATTAATCATTAACCCAAATTTTGAACCTGTATAATTTCTTAACTTATTCACATATTGT	648
Db	68	GTTTTTTGAAGGCTTATTTTTTTATTTTTTGTGTTTTTTTTTTTTTTTTTTTTTTT	12

RESULT 15

US-09-949-016-12386

; Sequence 12386, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12386

; LENGTH: 263693

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-12386

Query Match	8.5%; Score 84.2; DB 4; Length 263693;
Best Local Similarity	51.1%; Pred. No. 1e-06;
Matches 247; Conservative	0; Mismatches 233; Indels 3; Gaps 2;
QY	83 TAAATACATATTGCATTTATCTCATTTCAATTTATTTATTAATTTATTTATTTATTTGGAT 142
Db	251609 TCRAAAGCAGTATCTATATGTAATTTATATTTAATTAATAAATAATATATATCTTTAA 251668
QY	143 AAAAAATTCCTAATCTTTACTTTTTTTTAAAAAGAAATTTATTTAAATTTATTTATATATTT 202
Db	251669 ATAAAAATTATATAAATTTAAATTTAAATTTAATTTATATATAATTTAAATTTAAATTTATAT 251728
QY	203 AGATAAAATTCCTAATCTTTACTTTTTTTTTTAAAAAGAAATTTCAATTCGGTTTTTTCTT 262
Db	251729 ATGATTTAAATTTAAATTTAAATTTAATTTATATATAGATTT--AAATTTAAATTTAATTTATATGATTTA 251786
QY	263 AATTTAGTTTTTAAATTTCTATACFAAATTTATAAAAAATTTCTGATCGGATTAGTGCGGTGCATA 322
Db	251787 AITTTAAITTTAATTTATATATAATAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 251846
QY	323 GTCAGTCACATGAATTTTGGTGGAGAAAAATAAAAAATTTAAACACATTTTTTCGATTAAT 382
Db	251847 TTATATATAAATTTAAATTTAAATTTATATATAAATTTAAATTTAAATTTAAATTTAAATTTAA 251906
QY	383 TTATTTATATATATAATAATAAACAACATTTTTTATTTAATTTGTGCAATPAATTTTTTTTA 442
Db	251907 TTATATATAATTTAAATTTAAATTTAAATTTAATTTAATTTAATAATTTAAATTTGATTAAC 251966
QY	443 ATTTAAAT--TTCAGCAACAATTTACACCTCATCTCAATTTAAATTTAATCTTATTAACCATAA 501
Db	251967 ATTTAAATTTATATAAATTTTATTTAAACACTCRAAAATCAAGAAATGTAAAGTAGTAA 252026
QY	502 TTAAATTTGTGAGGACAAATTTATTTTTTAAATCTCACCTCCATTTAAATGCATATTTAATTT 561

Db 252027 ATCAAAGACCTTAATCAATACTGATCCTATTTCCTGCTCATACATAAGCATTATATTT 252086

Qy 562 TTT 564

|||

Db 252087 TTT 252089

Search completed: March 15, 2005, 02:00:37
Job time : 214 secs

THIS PAGE BLANK (USPTO)

```

293  ATTTTTTTAAAAATTATTTTTTTAAAAAATAAATAAAAAATAAATAATA 467
352  AAAATAAAAAATAAACACATTTTTTCGATTAATTTATTTATATATATAATAATAAACACAT 411
353  ATTTAAAAAATAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 412
412  TTTTATTTTAATGCT---TGTCAAATAATTTTTTTAAATTTAAAAATTCAGCACAAATTC 467
413  ATTTATTTAATTTATTTTATTTTTTTTAAAAATAATAAAAAAATAATTTTAAAAAATAA 472
468  ACTCTCATCATTAATTTTAAATCTTATTACCATATAATAAAAATTTGTGAGGACAAATTATTTT 527
473  AATATATTTTAAATAAAAAAATAAAAAATTTTATAATTTTATAATTTAAAAATTTATATAT 532
528  TAATCTCACCTCCATTAATGCATATTTATTAATTTTTTTTGGTCGATCTCTTATTTCACCTC 587
533  TAAAAATTAATAATTTTATATTAATTTATAACGTTTTTTATTAATAATTTATATTTTAATA 592
588  CTAACATTAATCATTTAAACCCAAATTTGAATCTGTTTATTAATTTCTTTAACTTAATTCCTA 644
593  ATAAATTTTATAAAAAAATAATTTATAAAATTTTAAAAATTTTTTAAAAAATAATTTTATA 649

RESULT 3
US-10-211-179-11/c
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTYROSYL PHOSPHATASE ACTIVITY
; FILE REFERENCE: PTS-0011
; CURRENT APPLICATION NUMBER: US/10/211,179
; NUMBER FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 119
; OTHER INFORMATION: n = a, t, c, or g
US-10-211-179-11

```

RESULT 2
US-10-473-126-240
: Sequence 240, Application US/10473126
: Publication No. US20040234973A1
: GENERAL INFORMATION:
: APPLICANT: Epigenomics AG
: TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cells
: TITLE OF INVENTION: proliferative disorders

```

: GENERAL INFORMATION:
: APPLICANT: Epigenomics AG
: TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
: proliferative disorders
: TITLE OF INVENTION:
:

```

[illegible]

US-10-473-126-386/c
; Sequence 386, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:

```

; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 386
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-386

```

Query Match	10.3%;	Score 101.6;	DB 18;	Length 8056;
Best Local Similarity	49.4%;	Pred. No. 8.3e-05;		
Matches 291;	Conservative 0;	Mismatches 296;	Indels 2;	Gaps 1;
Qy	53	ATTTTTTATTTTACCTCGCTAGATTGCTAAATCTATTGCAATTTATCTCATTTTCATT	112	
Db	2420	ATTTTTTTTTTAAATTTCAAAAAATAAAAAATTAATAATTTATTTATAAAAAATAAAAA	2361	
Qy	113	ATTTATTTAAATTTATTTATATTTATTTGGATAAAAAATCTAATACTTTTACCTTTTTTAAA	172	
Db	2360	ATTATATTTTAAAAATTAATAAAAAATTTATTTAAAAATTAATAAAAAATTTTATAA	2301	
Qy	173	AAGAAATTTATTTAAATTTATTTATTTATTTAGATAAAAAATCTAATACTTTTACTTTTTT	232	
Db	2300	AATAAATTAATAATAAAAAATTTAAAAATTTTAAATTTATTTTAAATATTTAAAAAATAAAAT	2841	
Qy	233	TTAAAAAGAAATTTCAAATTCGGTTTTTTTCTTAAATTTAGTTTTTAATCTATACTAATATAA	292	
Db	2240	TTTTATTAATAATTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTAA	2181	
Qy	293	AAATTCGATCGGAATAGTGTGGTGTCAAAGTCAAGTCACANGAATTTTGTGGAGAAAA	352	
Db	2180	TAAAAAT--ATAATTTTTTTTTTTTTTTTTTTTTTTTTTAAATAAAAAATAAAAAATTTATTA	2123	
Qy	353	AAATAAAATTTAACAACATTTTTTCGAATTAATTTATATATATATATAATAATATAAACACATT	412	
Db	2122	AAATATATTAATAAATAATTAATAATTAATTTATTAATAATTTTAAATAATTTTTTAATAAT	2063	
Qy	413	TTTTATTTAAATGTTGCAATAATATTTTTTTTAAATTTAAAAATTTTCAACACAACTTACACTCT	472	
Db	2062	TTTTTAAAAATAAAAAATAATTTTAAAAATAATTTTTTATTTTAAATATATTTAAAAATTTTAT	2003	
Qy	473	CATCATTTAAATTTAATCTTATTACATAATTTAAAAATTTGTGGAGCAATTTATTTTTTAATC	532	
Db	2002	TAAATTTTAAATAAAAAATTAATAATTAATTTTATTTTAAAAAATAATATATTTTTTTTTTATT	1943	
Qy	533	TCACCTCCATTAATGCAATATTATTAATTTTTTGTTCGATACCTCTTATTTTCACTCCCTAAC	592	
Db	1942	TATATTAATAATTAATTTATTTTTTTTTTTTTTATAAAAAATTTTATTTTATAATAAAAAAAT	1883	
Qy	593	ATTAATCATTTAACCCAAATTTTGAACCTGTTATTAATTTCTTAACCTATTATCA	641	
Db	1882	AAAAATAAATTAATAAATAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAA	1834	

RESULT 7
US-10-473-126-240/c
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 240

```

; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-240

```

Query Match	10.2%	Score 100;	DB 18;	Length 8056;
Best Local Similarity	49.2%;	Pred. No. 0.00014;		
Matches 290; Conservative	0;	Mismatches 297;	Indels	2; Gaps 1;
Qy	53	ATTTTTTATTTTACCTCCCTAGATTGTAATACTATTGCATTTATCTCATTTTCATT	112	
Db	2420	ATTTTTTTTTTAAATTTTCGAAAAATTAATAAAATTTTAATAATTTATTATATAAAATATAAAA	2361	
Qy	113	ATTTATTTTAAATTTATTTATATTATTGGATPAAAAATTTCTAAATCTTTTACTTTTTTTAA	172	
Db	2360	ATTATATTTAAAAATTTAAATAAAAAATTTATTTAATACGAAAAATTAATAAATTTATTTTAA	2301	
Qy	173	AAGAAATTTATTAATTTATTTTATATTATTAGATAAAAAATTCATAACTTTTACTTTTTTT	232	
Db	2300	AAATAAATAAAATTTAAAAATTTTAAATTTTATTATTTTAAATATATAAAAAATAAAT	2241	
Qy	233	TTAAAAAGAAATTTCAATTCGGTTTTTTCTTAAATTTAGTTTTTAATCTATACATAATTA	292	
Db	2240	TTTTATTAATAATTAATTAA	2181	
Qy	293	AAATTCGTGATCGGAATTAGTCGGTGTCAAAGTCAAGTCACATGAATTTTGTGGAGAAA	352	
Db	2180	TAAAAAT - ATAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTAAATAAAAAATTA	2123	
Qy	353	AAATAAAATTAACACATTTTTTCGATTAATTTATATATATATATAATAATATAAACA	412	
Db	2122	AAATATATTAAATAAATTTATAATTAATTTATTTATATAATTTTAAATATTTTAAATA	2063	
Qy	413	TTTTATTTAAATGTTGTCAAATAATATTTTTTTTAAATTTAAAAATTTTCAACACAAT	472	
Db	2062	TTTTTAAAAATAAAAAATATTTTAAAAATAATATTTTTTATTTTAAATATATTTAAAA	2003	
Qy	473	CATCATTTAAATTTAATCTTTATACCAATAATAAAATTTGTGGGACAATTTATTTTTTA	532	
Db	2002	TAAATTTTAAATAAAAAATTAATAATTTATTTTAAATAAAATTTATTTTTTTTTTA	1943	
Qy	533	TCACCTCCCAATTAATGCATATTTATTAATTTTTTGTTCGATACTTCTTATTTTCACTC	592	
Db	1942	TATATATAATTTAAATTTATTTTTTTTTTTTTTTTATAAAAATTTTATTTTATATAA	1883	
Qy	593	ATTAATCATTTAACCCAAATTTTGAACCTGTATAAATTTCTTAACCTATTCA	641	
Db	1882	AAAAATAAATTTTAAATTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAA	1834	

8 J. LISA

```

RES001.8
US-10-425-115-132399/c
? Sequence 132399, Application US/10425115
? Publication No. US20040214272A1
? GENERAL INFORMATION:
? APPLICANT: Thomas J.
? APPLICANT: La Rosa, Thomas J.
? APPLICANT: Kovalic, David K.
? APPLICANT: Zhou, Yihua
? APPLICANT: Cao, Yongwei
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: Plants
? FILE REFERENCE: 38-21(53222)B
? CURRENT APPLICATION NUMBER: US/10/425,115
? CURRENT FILING DATE: 2003-04-28
? NUMBER OF SEQ ID NOS: 369326
? SEQ ID NO 132399
? LENGTH: 1130
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: unsure

```

Query Match 9.8%; Score 96.2; DB 15; Length 15548;
Best Local Similarity 53.7%; Pred. No. 0.00054;
Matches 267; Conservative 0; Mismatches 223; Indels 7;

8272-CCF-TTC-OT-SO

Query Match 9.8%; Score 96.2; DB 15; Length 15548;
Best Local Similarity 53.7%; Pred. No. 0.00054;
Matches 267; Conservative 0; Mismatches 223; Indels 7;


```

; NAME/KEY: unsure
; LOCATION: (214, 2796...2797, 4347)
US-10-221-613-33

Query Match          9.7%; Score 95.4; DB 17; Length 6109;
Best Local Similarity 49.7%; Pred. No. 0.00055;
Matches 296; Conservative 0; Mismatches 296; Indels 3; Gaps 2;

QY 53 ATTTTATTTTACCTCGCCCTAGATTCGTAAATACTATGCAATTTATCTCATTTTCAAT 112
Db 1628 ATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1687

QY 113 ATTTATTTTAAATTTATTTATTTATTTGGATTAATAATCTAAATCTTACTTTTAA 172
Db 1688 ATATATCGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1747

QY 173 AAGAAATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 232
Db 1748 ATGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1807

QY 233 TTTAAAGAAATTTCAATTCGGTTTTCTTAAATTTAGTTTTTAAATCTTAAATTTATA 292
Db 1808 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1865

QY 293 AAATTCGATCGGAATTTAGTGGTGCAAGTCAAGTCAATTTTGGTGGAGAAA 352
Db 1866 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1925

QY 353 AATAAAATTAACACATTTTCGATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 412
Db 1926 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1984

QY 413 TTTATTTAATGTTGTCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 472
Db 1985 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2044

QY 473 CATCATTAATTTAAATCTTTATTAACCAATTAATAATTTAAATTTTAAATTTTAAATTTTAA 532
Db 2045 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2104

QY 533 TCACCTCCATTAATGCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 592
Db 2105 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2164

QY 593 ATTAATCATTAACCCCAATTTGAACTGTTATTAATTTCTTAACTTATTTCACTATTTG 647
Db 2165 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2219

RESULT 13
US-10-311-455-952
; Sequence 952, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 952
; LENGTH: 6963
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-952

Query Match          9.6%; Score 94.6; DB 15; Length 6963;
Best Local Similarity 49.9%; Pred. No. 0.00074;
Matches 290; Conservative 0; Mismatches 289; Indels 2; Gaps 2;

QY 53 ATTTTATTTTACCTCGCCCTAGATTCGTAAATACTATTTGCAATTTATCTCATTTTCAAT 112
Db 1957 AATTTTGAATTTTCGGTTTTTAGGTTTTTTTGAAGATTTTCTGTTGCTATAGGTTTTAAAT 2016

QY 113 ATTTATTTTAAATTTATTTATTTATTTTGGAT-AAAAATCTCAATCTTACTTTTAA 171
Db 2017 AGTGAATATTTAAATTTTATTTATTTTAAATGTTTTAGTTTTTATTTATTTATTTATTTATTT 2076

QY 172 AAAGAAATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 231
Db 2077 GTTTTTTTTAAAGATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2136

QY 232 TTTAAAGAAATTTCAATTCGGTTTTCTTAAATTTAGTTTTTAAATCTTAAATTTATA 291
Db 2137 TTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2196

QY 292 AAAATTCGATCGGAATTTAGTGGTGCAAGTCAAGTCAATTTTGGTGGAGAAA 351
Db 2197 ATTTTAGTAATATAGATTTTGGTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2256

QY 352 AATAAAATTTAAACACATTTTTCGATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 411
Db 2257 AGATAAAATTTAAAGTTTGTGTTTTTATTTATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 2316

QY 412 TTTTATTTTAAATGTTGTCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 471
Db 2317 TTTAAATTTAAATTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 2376

QY 472 TCATCATTAATTTTAAATCTTTATTAACCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 531
Db 2377 GTGGAATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2436

QY 532 CTCACCTCCAT-TAATGCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 590
Db 2437 ATAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2496

QY 591 ACATTAATCATTAACCCCAATTTTGAACCTGTTATTAATTTCTT 631
Db 2497 AATTTTAAATATGAATTAATTAATTTTAAATTTT 2537

RESULT 14
US-09-960-352-11234
; Sequence 11234, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.0067/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11234
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 48-LIB3058-052-Q1-K1-DB
US-09-960-352-11234

Query Match          9.6%; Score 94.4; DB 9; Length 419;
Best Local Similarity 52.6%; Pred. No. 0.00039;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 23:07:26 ; Search time 3170 Seconds
(without alignments)
11827.541 Million cell updates/sec

Title: US-10-089-543-2_COPY_449_1433

Perfect score: 985
Sequence: 1 gatgtgaccagtgttgag.....tgagagaaatccttcacatc 985

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	131.6	13.4	1101	9	CNS000EVL
C 2	120.2	12.2	1101	9	CNS003BD
C 3	120.2	12.2	1101	9	CNS00EVL
C 4	120	12.2	1352	9	AG381852
C 5	116	11.8	1200	9	CNS016CO
C 6	115	11.7	1441	9	AG396166
C 7	114.8	11.7	1275	9	CL033318
C 8	114.4	11.6	886	8	BH172777
C 9	114.4	11.6	886	9	CNS07JUX
C 10	114	11.6	1101	9	CNS00EO7
C 11	113	11.5	1542	9	AG386981
C 12	112.8	11.5	839	5	BQ151187
C 13	112	11.4	990	9	CNS006OI
C 14	111.8	11.4	1896	9	CG753083
C 15	111.6	11.3	810	9	AG486400
C 16	111.4	11.3	1029	9	AG381ZGM
C 17	111	11.3	1391	9	CG754863
C 18	111	11.3	1493	9	CL078589
C 19	110.8	11.2	1242	9	CL068807
C 20	110.6	11.2	932	9	CL479576
C 21	110.4	11.2	1378	9	AG350209
C 22	109.6	11.1	1348	9	CG749499
C 23	109.2	11.1	1172	3	CR734084
C 24	109.2	11.1	1392	9	CG757503

C 25	109	11.1	939	6	BY720774
C 26	109	11.1	1005	6	BY720771
C 27	108.8	11.0	1101	9	CNS003BB
C 28	108.8	11.0	1127	9	CL128786
C 29	108.8	11.0	1227	9	AG430010
C 30	108.6	11.0	1268	8	BZ577630
C 31	108.6	11.0	1608	9	CL118721
C 32	108.4	11.0	1015	9	CL135318
C 33	108.4	11.0	1568	9	AG321250
C 34	108.2	11.0	1228	9	CL104752
C 35	108.2	11.0	1291	9	AG430338
C 36	108	11.0	1101	9	CNS0039G
C 37	108	11.0	1224	9	CL077121
C 38	107.8	10.9	988	9	CL087333
C 39	107.6	10.9	1210	9	CG749728
C 40	107.4	10.9	1243	9	AG390983
C 41	107.2	10.9	1314	9	CL077082
C 42	107	10.9	1017	9	CL101048
C 43	107	10.9	1310	9	AG382239
C 44	106.8	10.8	1201	9	CNS0167M
C 45	106.6	10.8	674	7	CV117970

ALIGNMENTS

CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL069706
VERSION AL069706.1 GI:4949849

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooeogawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match 13.4%; Score 131.6; DB 9; Length 1101;
Best Local Similarity 35.6%; Pred. No. 7.5e-11;


```

fly), genomic survey sequence.
AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
          collaboration with the Berkeley Drosophila Genome Project (BDGP).
          The BDGP is constructing a physical map of the Drosophila
          melanogaster genome using these BACs. For further information
          please see http://www.fruitfly.org The BDGP Drosophila
          melanogaster BAC library was prepared by Kazutoyo Osoegawa and
          Aaron Mammoler in Pister de Jong's laboratory in the Department of
          Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
          NY. The library is named RPCL-98 and was constructed by partial
          EcoRI digestion of Drosophila DNA provided by the BDGP from the
          isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
          P1 and EST libraries. A more detailed description of the library
          and how to order individual BAC clones, the entire library, or
          filters for hybridization from the BACPAC Resource Center can be
          found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES             source
  1. .1101
     /organism="Drosophila melanogaster"
     /mol_type="genomic DNA"
     /db_xref="taxon:7227"
     /clone="BACR29B23"
     /clone_lib="RPCL-98"
     /note="end : T7"

ORIGIN
Query Match          12.2%; Score 120.2; DB 9; Length 1101;
Best Local Similarity 37.1%; Pred. No. 4.3e-09;
Matches 217; Conservative 113; Mismatches 253; Indels 2; Gaps 1;

QY 66 CCTCTGCTAGATTCGTAAATCTACTATGCAATTTATCTCATTTTCAATTTATTTTAAATTA 125
DB 434 CCCSCSCCMWMMHMAATYTCTCAHTWMMWMAATWMAAATAWAAWAAATTTATWAA 493

QY 126 TTTTATATATTTGGATATAAATCTTAATCTTTACTTTCTTTTAAAAAGAAATTTATTA 185
DB 494 TWAATAAAWMMWATTTTTTMMWMTWATTTTMMWMTWATTTAAATAAAWAAWATAATTT 553

QY 186 ATTATTTTATATTTAGTAAATCTTAATCTTTACTTTCTTTTAAAAAGAAATTT 245
DB 554 AAATAAATAWATAAATAATTTAAWAAWATAATTAATTAATTAATAATWATAATAAATA 613

QY 246 CAATTCGGTTTTTCTTAATTAATAGT--TTTAATCTTACTAATATAAAAAATTCGTATC 303
DB 614 ATATTTTATTAATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 673

QY 304 GGATTAAGTGGTGGTCAAGTCAAGTCAATGAATTTTGTGGAGAAAAATAAATAATTA 363
DB 674 TAAATAATTTTAAATAWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 733

QY 364 AACACATTTTTCATTAATTTATATATATATATATATATATATATATATATATATAT 423
DB 734 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 793

QY 424 TTGCTCAATAATATTTTAAATTAATAATTTTCAAGCAACAATCACTCATCATTAAT 483
DB 794 ATAAWAAWAAWATAAATAWATAWATAWAAWAAWATAWATAWATAWATAWATAWATAW 853

QY 484 TTAATCTTATACCAATAATTAATAATTTGTGGAGCAATATTTTAAATCTCACCCCTCAT 543

```

```

DB 854 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 913
QY 544 TAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 603
DB 914 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 973
QY 604 ACCCAATTTGAACCTGTTATATAATTTCTTAACATTTTACATTTATTTGT 648
DB 974 TWTATAWTTTATATATATATATATATATATATATATATATATATATATATATTA 1018

RESULT 4
AG381852
LOCUS Mus musculus molossinus DNA, clone:MSMg01-192A15.T7, genomic survey
DEFINITION sequence.
ACCESSION AG381852
VERSION AG381852.1 GI:47993057
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1352)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:hattori@gsc.riken.jp. URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
          library availability, please contact Kunya Abe (abe@rtc.riken.jp).
          Tsukuba Institute, Bio Resource Center,
          The Institute of Physical and Chemical Research (RIKEN) 3-1-1
          Koyadai, Tsukuba, 305-0074 Japan
          phone: 81-298-36-9189, fax: 81-298-36-9199
          e-mail: abe@rtc.riken.jp
          PRIMERS
          Sequencing : T7
          LIBRARY Vector : pBACe3.6
          R.Site 1 : EcoRI.
          R.Site 2 : EcoRI.
FEATURES             source
  1. .1352
     /organism="Mus musculus molossinus"
     /mol_type="genomic DNA"
     /sub_species="molossinus"
     /db_xref="taxon:57486"
     /clone="MSMg01-192A15.T7"
     /sex="male"
     /tissue_type="mixture of kidney and spleen"
     /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match          12.2%; Score 120; DB 9; Length 1352;
Best Local Similarity 50.0%; Pred. No. 4.5e-09;
Matches 294; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 52 AATTTTATTTTACCTCTGCTAGATTCGTAATAACTATTCGATTTATCTCATTTTCAT 111
DB 727 ATTTTATTAATTTTATTTATTTTATTTTATTAATAATATTTTATTTTATTTTATTTT 786

QY 112 TATTTTATTAATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 171
DB 787 TTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 846

```

```
QY 172 AAAGAAATTTAATTAATTTATATATTTATATTTAGATAAATAATCTAATCTACTTACTTTT 231
    |||||
Db 847 TTTNTTTTATTTTGAATAATTTTATATATTTATATTTAGTTTATTTATTTTATTTT 906
    |||||
QY 232 TTTAAAGAAATTTCAATTCGGTTTCTTAATTTAGTTTAAATTTCTATATACTAATTA 291
    |||||
Db 907 ATTATTTATATATTTAATATATATTAATTTATTTTAAATTTTAAATTTAAATTTAT 966
    |||||
QY 292 AAAATTTCTGATCGGATAGTGGTGCAGAGTCAAGTCAATGACATGAAATTTTGGTGAGAAA 351
    |||||
Db 967 AAAATTTTATATATTTTATTTTATTTTATATTTTATTTTATTTTATTTTATTTAT 1026
    |||||
QY 352 AAATAAAATTAACACATTTTCGATTAATTTATATATATATATATATAATAAACAAT 411
    |||||
Db 1027 ATTATTAATTTATTTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 1086
    |||||
QY 412 TTTTATTTAATGTTGTCATATATATTTTAAATTTAAATTTTCAAGCACAAATTTACACTC 471
    |||||
Db 1087 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1146
    |||||
QY 472 TCATCAATTAATTAATTTTATTAACATTAATTAATTTTGGTGAGGCAATTTTAAAT 531
    |||||
Db 1147 ATTATTTTATTTATATATATATATATTTTATTTTATTTTATTTTATTTTATTTTATTT 1206
    |||||
QY 532 CTACCTCAATTAATGCAATTTTAAATTTTGGTTCGATCTCTTATTTTCACTCCTAA 591
    |||||
Db 1207 TTCTCTCTATTTTATCCCTTTATTTCTTTTCTTTTCTTATATCTTCTCTATTTACTT 1266
    |||||
QY 592 CATTAAATCAATTAACCAATTTTGAACCTGTTTATATTTTCTTAACTTAT 639
    |||||
Db 1267 CATCACTATCTACTCTATTTTCTCTCTTATATTTCTTTCTTATTTCTTCTTATTT 1314
    |||||

RESULT 5
CNS016CO 1200 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence 17 end of BAC
DEFINITION BACN15E04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL106578
VERSION AL106578.1 GI:5622626
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1200)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES
    source
    1..1200
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BACN15E04"
    /clone_lib="DrosBAC"
    /plasmid="pBelobAC11"
    /note="end : 17"

ORIGIN
Query Match 11.8%; Score 116; DB 9; Length 1200;
```

```
Best Local Similarity 39.6%; Pred. No. 1.9e-08;
Matches 164; Conservative 86; Mismatches 164; Indels 0; Gaps 0;

QY 114 TTTATTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 173
    |||||
Db 1200 TATATTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1141
    |||||
QY 174 AGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 233
    |||||
Db 1140 AWAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1081
    |||||
QY 234 TAAAAAGAAATTTCAATTCGGTTTCTTAATTTAGTTTAAATTTCTATATACTAATTA 293
    |||||
Db 1080 TTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1021
    |||||
QY 294 AATTTCTGATCGGATAGTGGTGCAGAGTCAAGTCAATGACATGAAATTTTGGTGAGAAA 353
    |||||
Db 1020 TATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 961
    |||||
QY 354 ATAAAAATTTAAACACATTTTTCGATTAATTTATTTATTTATTTATTTATTTATTTATTT 413
    |||||
Db 960 TTAATTTWAAAAAAWATTTTWTTRTTWWAAATTAATTTTAAAAATTTTTCGTTAAWATTA 901
    |||||
QY 414 TTATTTAATGTTGTCATATATTTTAAATTTTAAATTTTCAAGCACAAATTTACACTCTC 473
    |||||
Db 900 WTCHTTATTTTWTARAAAAATTTTAAATTTTATTTATTTTATTTATTTATTTATTTATTT 841
    |||||
QY 474 ATCATTTAAATTTAATCTTATTACCATTAATTTAAAAATTTGTGAGGCAATTTATTTT 527
    |||||
Db 840 AACTTTWAGHAARGCYYYANCCCKKKKKKKTTTTTTTGGSGGGGDKKTTY 787
    |||||

RESULT 6
AG396166 1441 bp DNA linear GSS 03-JUN-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-216D06.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG396166
VERSION AG396166.1 GI:48029247
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1441)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan 1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACE3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
Location/Qualifiers
    1..1441
    /organism="Mus musculus molossinus"
    /mol_type="genomic DNA"

FEATURES
    source
```

/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-216D06.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"

ORIGIN

Query Match 11.7%; Score 115; DB 9; Length 1441;
Best Local Similarity 50.5%; Pred. No. 2.6e-08;
Matches 280; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 79 TTCGTAATAAATCTCAATCTACTTCTCTTTTAAAGAAATTTCAATTCGCTTTT 258
DB 639 TTTTATTTTAAATATTTATTTTATTTTATTTTATTTTAAATTTTATTTATTTAT 698
QY 139 GGATAAAATCTCAATCTACTTCTCTTTTAAAGAAATTTCAATTTATTTATTT 198
DB 699 ATTCTTTTATTTTATTTTGTATATTTTATTTTATTTTATTTTAAATTTTATTA 758
QY 199 ATTAGATAAATCTCAATCTACTTCTCTTTTAAAGAAATTTCAATTCGCTTTT 258
DB 759 ATTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 818
QY 259 TCTTAATTTAGTTTAAATCTACTAATTAATAAATCTGATCGGATGCTGCT 318
DB 819 TTTTATTTTATTTTATTTTAAATATTTATATTTTATTTTATTTTATTTTATTT 878
QY 319 CAAAGTCAATGTCATGAAATTTTGGAGAAAAATTAATAACACATTTTTCGAT 378
DB 879 TTTTATTTTATTTTAAATATTTTATTTTATTTTATTTTATTTTATTTTATTTA 938
QY 379 TAATTTATTTATTAATAATAATAACACATTTTATTTATTTATTTATTTTCAAT 438
DB 939 TTTTATTTTAAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 998
QY 439 TTTAATAAATTTGAGCAACAATACATCTCATATTAATAATTAATTAATTAACCA 498
DB 999 TCTTTTATTTATTTATTAATTTATATATTTATTTATTTTATTTTATTTTAT 1058
QY 499 TAATTAATAATTTGAGGACAAATTTTATTTTAAATCTCACCTCCATTAATGCA 558
DB 1059 TTATTAATAATTTATTTTATTTTATTTTATTTTAAATATTAATTTTATTTTCA 1118
QY 559 ATTTTGTTCGATCTCTTATTTCTACTCTCAATTAATCAATTAACCAATTTTCA 618
DB 1119 TATATTTTATTTTAAATTAATTAATTTATTTATTTATTTATTTATTTATTT 1178
QY 619 GTTATAATTTCTTAA 633
DB 1179 ATTTTATTTTAA 1193

RESULT 7
LOCUS BH177277/c
DEFINITION 1275 bp DNA linear GSS 31-DEC-2003
CH216-36F4 Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-36F4,
genomic survey sequence.
ACCESSION CL033318
VERSION CL033318.1 GI:40483882
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1275)
Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center

Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 40
High quality sequence stop: 194.
Location/Qualifiers
1. .1275
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-36F4"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

Query Match 11.7%; Score 114.8; DB 9; Length 1275;
Best Local Similarity 50.3%; Pred. No. 2.9e-08;
Matches 272; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

QY 108 TCATTATTTATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 167
DB 844 TTAATAATAAATTAATAAATAATTAATAAATAATTAATAAATAATTAATAAATAATTA 785
QY 168 TTAAGAAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 227
DB 784 TTTTAAATATTTATTTAAATATTTAAATATTTAAATATTTAAATATTTATTTATTT 725
QY 228 TTTTAAAGAAATTTCAATTCGCTTTTCTTTAAATTTAGTTTAAATTTCTATCTAAT 287
DB 724 TATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 665
QY 288 TATAAATAATTCGATCGGATTTAGTGGTGTCAAGTCAAGTCAAGTCAAGTCAAGT 347
DB 664 ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 605
QY 348 GAAAAATAAATAATTAACACATTTTTCGATTTATTTATTTATTTATTTATTTATTTA 407
DB 604 AATAAATAAATAATTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTA 545
QY 408 ACAATTTTATTTAATTTGTCATTAATTTTAAATTTAAATTTTAAATTTTAAATTTA 467
DB 544 ATAAATTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATA 485
QY 468 ACTCTCATCATTTAATTTAATCTTATTTACCAATTTAAATTTGTTGAGCAATTTAT 527
DB 484 TAAATTTTGTAAATTAATAATTTTATTTATTTATTTATTTATTTATTTATTTATA 425
QY 528 TAATCTCACCTCCATTAATGCAATTTATTTATTTATTTATTTATTTATTTATTTAC 587
DB 424 AAAAAATAAATAAATAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 365
QY 588 CTAAACATTAATCATTAACCAATTTTGAAGTCTGTTAAATTTCTTAACTTTATTC 647
DB 364 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 305
QY 648 T 648
DB 304 T 304

RESULT 8
LOCUS BH177277/c
DEFINITION 886 bp DNA linear GSS 19-OCT-2001
008 L 22-rev SmbAC1 Schistosoma mansoni genomic clone 008L22 5',
genomic survey sequence.
ACCESSION BH177277
VERSION BH177277.1 GI:16275873

[illegible]

COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS
Sequencing : TJ
LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

FEATURES Location/Qualifiers

1..1542

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMg01-201G10.TJ"

/sex="male"

/tissue_type="mixture of kidney and spleen"

/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 11.5%; Score 113; DB 9; Length 1542;

Best Local Similarity 49.3%; Pred. No. 5.3e-08; Indels 0; Gaps 0; Matches 290; Conservative 0; Mismatches 298;

QY 52 AATTTTATTTTACCTCCCTCGCTAGATTCGTAATCTGTAATCTATTCGATTTATCTATTCATTTCAAT 111
DB |||||
QY 324 ATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 383
DB |||||
QY 112 TATTTATTAATTTATTTATTTATTTATTTGGTAAATTTCTAATCTTTTACTTTTCTTTTAA 171
DB |||||
QY 384 TATATATTTTAAATATAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 443
DB |||||
QY 172 AAGAAATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 231
DB |||||
QY 444 TTTTATAAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 503
QY 232 TTTTAAAGAAATTTCAATTCGGTTTCTTAAATTTAGTTTATTTTAAATTTCTATCTATCTAATTTAA 291
DB |||||
QY 504 TATATTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATATA 563
QY 292 AAAATTCGATCGATTTAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 351
DB |||||
QY 564 ATATATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 623
QY 352 AAATAAATAATTTTAAACACATTTTTCGATTTTATTTTATTTATATATATATATATATATATATATATATATAT 411
DB |||||
QY 624 TTTTATTTTATTTTATTTTAAATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 683
QY 412 TTTTATTTTAAATTTGTCGATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 471
DB |||||
QY 684 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 743
QY 472 TCATCATTTAAATTTTAACTTTATTTACCATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 531
DB |||||
QY 744 TTAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 803
QY 532 CTCACCTCCATTAATGATATTTAATTTTCTGTTGATTTTCTGTTGATTTTCTGTTGATTTTCTGTTGATTTTCTG 591
DB |||||
QY 804 ATATATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 863
QY 592 CATTAATCAATTAACCAATTTTGACGTTTATATTTTCTTTAACTTTAT 639
DB |||||
QY 864 ATATATTTTATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 911

RESULT 12
BQ151187 839 bp mRNA linear EST 24-APR-2002
LOCUS BQ151187

DEFINITION NF048A07LF1050 Developing leaf Medicago truncatula cDNA clone
NF048A07LF 5', mRNA sequence.

ACCESSION BQ151187

VERSION BQ151187.1 GI:20288246

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM

Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 839)

AUTHORS

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Imman, J.T., Wellier, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula leaf-library

Unpublished (2000)

CONTACT: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 224 6650

Fax: 580 224 6692

Email: gdmay@noble.org

Insert Length: 839 Std Error: 0.00

Plate: 048 row: A column: 07

Seq primer: TCACACGAGAAACAGCTATGAC.

FEATURES

Location/Qualifiers

1..839

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="NF048A07LF"

/tissue_type="leaf"

/dev_stage="Pooled developmental"

/clone_lib="Developing leaf"

/note="Vector: Lambda Zap; Contains a mixture of very

young, developing, mature and senescing leaves."

ORIGIN

Query Match 11.5%; Score 112.8; DB 5; Length 839;

Best Local Similarity 50.7%; Pred. No. 6.3e-08;

Matches 298; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

QY 54 TTTTATTTTACCTCCCTCGCTAGATTCGTAATCTATTCGATTTATCTATTCATTTCAATTA 113
DB |||||
QY 174 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 233
DB |||||
QY 114 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 173
DB |||||
QY 234 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 293
QY 174 AGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 233
DB |||||
QY 294 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 353
QY 234 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 293
DB |||||
QY 354 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 413
QY 294 AATTCGATCGGATTTAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 353
DB |||||
QY 414 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 473
QY 354 ATAAATTTAAACACATTTTTCGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 413
DB |||||
QY 474 ATAAATTTATATATCCATAATATTTTATTTATTTATTTTATTTATTTTATTTATTTATTTATTTATTTATTTAA 533
QY 414 TTTTATTTATTTGTCATATATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 473
DB |||||
QY 534 TTTTATTTATATAATAATAAATAAATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 593
QY 474 ATCAATTTAAATTTAATCTTTATTTACCATAATTTAAATTTGTGGGACAATTTATTTTATTTATCT 533

